



UNIVERSITAT DE BARCELONA

Neural signatures of word learning in adults and children: a multi-methodological approach

Neus Ramos Escobar

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (www.tdx.cat) i a través del Dipòsit Digital de la UB (diposit.ub.edu) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tdx.cat) y a través del Repositorio Digital de la UB (diposit.ub.edu) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (www.tdx.cat) service and by the UB Digital Repository (diposit.ub.edu) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.



**Cognition and Brain
Plasticity Unit**



**UNIVERSITAT DE
BARCELONA**

Neural signatures of word learning in adults and children: a multi-methodological approach

Neus Ramos Escobar

Cognition and Brain Plasticity Unit

Department of Cognition, Development and Educational Psychology

University of Barcelona

Doctoral program in Brain, Cognition, and Behaviour

Supervisors:

Dr. Antoni Rodríguez Fornells

Dr. Clément François

Acknowledgements

Al llarg d'aquests anys de doctorat, especialment durant els últims anys, m'he trobat en moltes ocasions pensant en tota la gent a la que estic tan agraïda per haver-me ajudat d'una manera o altra en el llarg camí que ha sigut aquesta tesis. I ara, que em dispenso a posar-ho en paraules em passa com cada vegada que m'enfronto a un full en blanc: mil idees em venen al cap i no ser per on començar.

M'ha semblat més fàcil començar pel principi de tot, bastant abans de començar el doctorat en aquella època de la universitat en la que tens la pressió dels exàmens i les emocions es viuen a flor de pell. En aquella època vaig tenir la gran sort de topar-me amb un dels millors professors que he tingut mai i que m'ha servit d'inspiració ara que jo m'he convertit en professora. El **doctor Josep Garcia Valero**, que a part de ser un gran docent va saber donar-me la força i fer-me recuperar la passió per la ciència en un moment que l'havia perdut. Sense ell segurament no hauria seguit el camí de la ciència i no seria on soc ara.

Ara, si que començo pels meus inicis a Brainvitge, en aquella primera vegada que vaig entrar al despatx de l'**Antoni Rodríguez Fornells** de la 4a planta, i em va encantar trobar un despatx ple de llibres per tot arreu! Després d'una breu entrevista en Toni va accedir a que fes les pràctiques de grau amb ell (jo estava saltant d'alegria!). Qui hagués imaginat que aquella decisió provocaria un punt d'inflexió tant gran en la meua vida tant a nivell personal com professional. Gràcies Toni per donar-me la oportunitat i creure en mi. Al llarg d'aquest 7 anys m'has ajudat a créixer personal i professionalment. Gràcies per ajudar-me a descobrir el món de la neurociència i de la ciència en general. Gràcies per despertar la meua curiositat científica, per ensenyar-me que a vegades cal buscar i rebuscar per obtenir respostes. Gràcies per acompanyar-me en aquest camí i gràcies per tot el que m'has ensenyat.

El primer projecte en el que vaig col·laborar (i probablement el més bonic de tots) va ser a l'Hospital de l'Esperança. Vaig aprendre moltes coses col·laborant amb la **Jenny**: a ser rigorosa treballant, a com tractar els pacients, a escriure de forma científica, etc. Però allí mateix a l'hospital de l'Esperança també vaig tenir la sort de treballar amb persones meravelloses: la **Nohora i la Susana**, sempre disposades a ajudar-me en tot i a la doctora **Esther Duarte** sempre amb ganes de col·laborar i tirar endavant el projecte. Tinc molt bon record de l'etapa a l'Esperança i es gràcies a les persones de les que vaig estar envoltada: totes vosaltres i també

els pacients. Em considero afortunada d'haver pogut afegir el meu gra de sorra en la recuperació dels pacients.

Quan vaig acabar el màster va sorgir la oportunitat de començar un projecte de llenguatge en col·laboració amb el grup de Finlàndia liderat per en **Matti Laine**. Thanks, Matti, for your help at the distance: giving me very useful feedbacks about the manuscripts and the project results. It has been a great pleasure to work with you.

En este proyecto tuve la oportunidad de colaborar por primera vez en el que se convertiría en mi otro supervisor de tesis, el doctor **Clément François** (después de todos estos años, aun cuando digo tu nombre en voz alta oigo a Pablo en mi cabeza diciendo "Clement Fransua"). Clem, podría escribir una tesis entera solo con todas las cosas por las que te estoy agradecida... Gracias por enseñarme a ser rigurosa científicamente, a escribir correctamente, a enseñarme mil cosas sobre EEG, sobre ciencia y sobre la vida. Eres un gran compañero, un gran supervisor y un gran amigo. Gracias por todos los consejos, por todas las conversaciones que hemos tenido y por todos los viajes y aventuras que hemos vivido juntos. Una de las muchas cosas buenas que me llevo de esta tesis es haberte conocido y tenerte como amigo. Una vez más, ¡mil gracias por todo!

Hablando de viajes, me ha venido a la cabeza los viajes a Bélgica y a Marsella y el grupito de personas que han dejado huella en mí, en durante mi paso por Brainvitge: Ane, Joan, Clem y David Cucurell. ¡¡Gracias chicos por hacer esos viajes divertidísimos y llenos de risas y aventuras!!

Ane, coincidimos en el lab cuando tu terminabas y yo empezaba la tesis, pero eso no impidió que se forjara una super bonita amistad entre nosotras que ha hecho aún más grande con el tiempo y la distancia. Nosotras, como dice Pereza, no queremos ser como los demás y especialmente tu para mi tienes una personalidad única. Eres autentica y natural, sin vergüenza a nada y desprendes una energía positiva que es adictiva y hace que sea imposible estar triste o enfadado estando a tu lado. No cambies nunca, sin duda eres otra de las cosas maravillosas que me ha traído esta tesis y mi estancia en Brainvitge.

Joan, estar al teu costat es sinònim de riures i converses interessants! Amb tu es impossible avorrir-se, tens una energia contagiosa! Gràcies per totes les converses que hem tingut, els riures, els dinars i els viatges junts (com el dia de la neu!). Gràcies per estar sempre disposat a ajudar-me. Admiro molt la teva eficiència professional i totes les coses que saps fer!

David Cucurell, encara recordo els primers dies a Brainvitge quan vaig iniciar-me amb el Matlab, tu vas ensenyar-me les primeres passes i vas fer que cada cop sigues més autònoma. Gràcies per la teva ajuda en els moments de crisi de registres i també pels teus consells quan estava davant una situació difícil. Espero que la nostra amistat segueixi molts anys (a pesar del confinament! 😊).

L'**Helena** també ha sigut un pilar clau en la meva estància a Brainvitge. Les nostres converses i ànims mutus han estat vitals per treure forces per seguir endavant. Helena, admiro de tu lo bona persona que ets i com et preocupes pels altres. Gràcies per la teva companyia i suport aquests anys.

L'altre pilar emocional sobretot en la recta final de la tesis es l'**Alba**. Gràcies Alba per les nostres converses donant-nos ànims per seguir i les tardes de companyia despotricant de la ciència i de la vida... Ha sigut un plaer compartir aquest viatge amb tu.

Després hi ha el **Pablo i la Lucía**, dos amics que tot i la distància també m'emporto d'aquesta tesi. Pablo t'admiro molt com a professional i com a persona, sempre treus un moment per ajudar als altres i ets sempre l'ànima de la festa. Gràcies per ajudar-me amb l'fMRI, per acollir-nos a casa teva i gràcies per tots els bons moments que ens has fet passar! Estar amb tu es sinònim de passar-ho bé! Y a ti **Lucía**, gracias por tus consejos, tu ayuda y por los buenos ratos que hemos pasado juntas. ¡Tienes un corazón enorme!

Però la família de Brainvitge no s'acaba aquí i son moltes més les persones que m'han ajudat i han estat al meu costat durant aquest camí. L'**Emma**, companya d'aventures en el projecte de les escoles, gràcies per la teva ajuda i el teu suport. Ets una gran persona i científica i arribaràs allí on et proposis! També el **Guillem i la Helena Tajadura** que m'han ajudat moltíssim els dies de registres a les escoles, gràcies als dos per la vostra ajuda! Gràcies també al **David Morales** per ajudar-me amb els experiments i en la revisió de l'anglès d'aquesta tesis al començament de tot.

També vull fer un especial agraïment als altres PI del grup: la **Ruth, l'Estela, el Josep i el Lluís**. Moltes gràcies per fer que Brainvitge tingui aquest esperit de companyerisme i que l'essència de Brainvitge es mantingui. Sense vosaltres això no hagués estat possible. També vull agrair a la **Anna Martinez** per la seva energia contagiadora i pel seu suport en la recta final de la tesis. A totes les noies de Brainvitge: la **Claudia, la Berta, la Patricia, la Gemma, la Audrey** i a tota la gent de Brainvitge que m'han acompanyat durant aquest temps.

Another very special time of the thesis has been the time I have spent in Marseille. I couldn't have chosen a better place to go. Thank you, **Daniele**, for greeted me warmly, thank you for teaching me with patience all that I've learned from you. It has been an absolute pleasure to meet you and to work with you. **Celine**, thank you for taking me in your place, thank you for all the good moments we have shared (the cycling trip, the nights with cheese and wine talking, the boat trip...). The "punaises du lit" took me the opportunity to share more time with you! One of the best things of my stay in Marseille is to know you and having you as a friend! I hope our friendship lasts for years!! And thank you to all the team guys: **Jacques, Arnaud, Jeremy, Manuel, Ben & Patrick**. Thank you for making my stay so grateful, and thanks for the great time we spent at the mountains! You're so boring, guys! =)

I ara arriba el moment d'agrair a tota la resta de persones de casa que m'han donat suport aquests anys. Com tampoc sé per on començar començaré per la més gran de la família, la nostra **Yaya**. Yaya, ets una font d'inspiració per a mi. Has estat donant-nos suport tota la vida, ens has ensenyat a l'Alicia i a mi una infinitat de coses. Hem passat moments inoblidables a Cubelles. Gràcies per ensenyar-me a ser treballadora, a ser estalviadora, constant i gràcies per transmetre'm la teva afició i passió per la cuina! Gran part del que soc ara es gràcies al que m'has ensenyat! Gràcies per tot! No has de patir tant per nosaltres, gràcies a tot el que ens has ensenyat estem preparades per a qualsevol cosa!!

Mama gràcies per tot el teu suport al llarg d'aquest anys. Gràcies per estar sempre al meu costat donant-me suport quan no podia més o quan les coses se'm feien costa a dalt. Admiro molt de tu la capacitat que tens d'aguantar i superar tot el que et posen davant (espero haver heretat una mica d'aquesta qualitat teva). He après de tu a ser treballadora i a lluitar pels meus objectius. Gràcies per fer-me veure que ho podia aconseguir. Gràcies per estar sempre allí.

Papa muchas gracias por todas las cosas que nos has enseñado y hemos aprendido de ti: a valernos por nosotras mismas, a hacernos ver que no necesitamos a nadie más que a nosotras para poder con todo lo que nos pongan por delante, a ser luchadoras y a ser trabajadoras. Gracias por todos los esfuerzos que habéis hecho tú y la mama para darnos lo mejor y habernos ayudado a llegar donde estamos ahora.

¡Y mi alma gemela, **Alicia**! Companya de vida i d'aventures. Gràcies per estar sempre al meu costat, per defensar-me davant de qualsevol cosa, per recolzar-me sempre, per animar-me i per estar sempre allí. Ets un pilar tant important a la meva vida que no se que faria sense tu. La teva

energia i vitalitat son contagioses, estar al teu costat es sinònim d'aventures i antònim d'avorrimet. T'estimo infinit.

Un dels altres pilars a part de la família que he tingut al llarg de la tesis (i de molt abans) son els meus estimats MIRANDEROS!! Les nostres "Cenas de finde" o millor dit "tech lessons de finde" em donen energia per seguir la setmana. Des de la època de la uni quan fèiem sopars a casa la Cris al sortir de la biblio, que després van passar a ser a casa de l'Orion, fins els sopars d'ara a casa l'Alicia (sense oblidar els nostres sopars virtuals cada dissabte durant el confinament). Sou un pilar essencial en la meva vida i em considero molt afortunada de tenir-vos com amics. **Laura, Àlvaro, Alicia, Georgina, Joan, Uri, Alba, Cris, Sergi, Alicia, Carlos, Jarek, Irene i Orion.** Gràcies per estar sempre!

Y una persona del grupo que se merece una especial mención. **Georgina**, gracias por estar siempre allí, por apoyarme siempre, animar-me y por estar a mi lado pase lo que pase!!

També vull agrair a l'**Aleix Puig** per la seva amistat tots aquests anys, per les hores a la biblio que hem passat junts i per sempre trobar una estona per veure'ns o parlar ara que estem lluny. Gràcies pel teu suport i la teva amistat tot i la distància que ens separa.

I he dejado para el final a esta persona tan especial al que pronto llamaré marido (aunque ya debería serlo...). **Edu**, muchas gracias por estar siempre a mi lado, por tu apoyo incondicional. Gracias por darme fuerzas cuando ya no me quedaban, gracias por darme ánimos cuando no podía más, gracias por hacerme reír cuando estaba llorando. Gracias por enseñarme a ver la vida con una sonrisa, por enseñarme que la vida es mejor si no te preocupas tanto por las cosas y enseñarme a disfrutar cada minuto que tenemos. Tu positivismo y alegría son contagiosos. Gracias por estar siempre a mi lado, no podía haber elegido a nadie mejor para compartir mi vida y me alegro que tú me hayas elegido a mí! =)

Index

| | |
|--|-----------|
| List of abbreviations | 1 |
| Abstract | 3 |
| Resumen | 5 |
| Chapter 1 Introduction..... | 11 |
| 1.1 General Introduction | 11 |
| 1.1.1 Neural bases of language | 14 |
| 1.1.2 Neural bases of language learning | 16 |
| 1.2 Language learning mechanisms | 17 |
| 1.2.1 Word segmentation and statistical learning..... | 17 |
| 1.2.2 Word learning and vocabulary acquisition: the world-to-word mapping problem | 26 |
| 1.3 Electrophysiological correlates of word learning | 36 |
| 1.3.1 Event-related potentials..... | 36 |
| 1.3.2 Brain oscillations | 50 |
| Chapter 2 Research aims | 63 |
| 2.1 Study I: The interplay between domain-general and domain-specific mechanisms during novel word learning | 63 |
| 2.2 Study II: Oscillatory activity and functional connectivity of concurrent word segmentation and meaning mapping in 9-year-old children | 64 |
| 2.3 Study III: The role of the hippocampus in SL | 65 |
| Chapter 3 Study 1 | 69 |
| 3.1 Introduction | 70 |
| 3.2 Materials and methods | 72 |
| 3.2.1 Participants..... | 72 |
| 3.2.2 Stimuli..... | 72 |
| 3.2.3 Procedure | 73 |
| 3.2.4 EEG data acquisition and analyses | 76 |

| | | |
|------------------|---|------------|
| 3.2.5 | Behavioral data analysis..... | 77 |
| 3.2.6 | ERP analyses..... | 77 |
| 3.3 | Results | 79 |
| 3.3.1 | Experiment 1..... | 79 |
| 3.3.2 | Experiment 2..... | 83 |
| 3.4 | Discussion | 85 |
| 3.5 | Conclusions | 88 |
| Chapter 4 | Study 2 | 93 |
| 4.1 | Introduction | 94 |
| 4.2 | Methods | 96 |
| 4.2.1 | Participants..... | 96 |
| 4.2.2 | Stimuli..... | 96 |
| 4.2.3 | Experimental Procedure..... | 97 |
| 4.2.4 | EEG data acquisition and analyses | 99 |
| 4.3 | Results | 100 |
| 4.3.1 | Behavioral results..... | 100 |
| 4.3.2 | Electrophysiological results..... | 101 |
| 4.4 | Discussion | 104 |
| 4.5 | Conclusion | 107 |
| Chapter 5 | Study 3 | 111 |
| 5.1 | Introduction | 112 |
| 5.2 | Methods | 114 |
| 5.2.1 | Participants..... | 114 |
| 5.2.2 | Data acquisition & electrode localization | 114 |
| 5.2.3 | Experimental procedure..... | 115 |
| 5.2.4 | Stimuli..... | 116 |
| 5.2.5 | SEEG Data analysis: Frequency tagging..... | 116 |
| 5.2.6 | Statistical analysis..... | 117 |
| 5.3 | Results | 118 |
| 5.3.1 | Behavioural results..... | 118 |

| | | |
|------------------|---|------------|
| 5.3.2 | Electrophysiological results..... | 118 |
| 5.4 | Discussion | 119 |
| 5.5 | Conclusions | 122 |
| Chapter 6 | General Discussion | 127 |
| 6.1 | Summary of results | 127 |
| 6.2 | Domain-general vs. domain-specific mechanisms | 129 |
| 6.3 | Brain regions involved in language learning | 131 |
| 6.4 | Simultaneous speech segmentation and word-referent mapping | 133 |
| 6.5 | Neuronal synchronization as a mechanism for language learning | 136 |
| 6.6 | Limitations and future directions | 139 |
| 6.7 | Conclusions | 142 |
| Chapter 7 | References | 147 |

List of abbreviations

| | |
|--------|---------------------------------------|
| AF | Arcuate Fasciculus |
| AFE | Ancient Farming Equipment |
| CFC | Cross-frequency coupling |
| Cons | Consistent learning group |
| CSC | Current source density |
| DP | Declarative/procedural model |
| DTI | Diffusion tensor imaging |
| EEG | Electroencephalography |
| ERP | Event-related potentials |
| FFT | Fast fourier transform |
| fMRI | Functional Magnetic resonance imaging |
| FN400 | Frontal N400-like component |
| ICA | Independent component analysis |
| IFG | Inferior frontal gyrus |
| IFOF | Inferior fronto-occipital fasciculus |
| ILF | Inferior longitudinal fasciculus |
| Incons | Inconsistent learning group |
| ITC | Inter-trial coherence |
| ITG | Inferior temporal gyrus |
| ITPC | Inter trial phase coherence |

| | |
|------|---|
| L1 | First language |
| L2 | Second language |
| LPC | Late positive component |
| MEG | Magnetoencephalography |
| MRI | Magnetic resonance Imaging |
| MTL | Medial temporal lobe |
| NIRS | Near-infrared spectroscopy |
| PC | Predictive coding |
| PLI | Phase lag index |
| PMC | Premotor cortex |
| sEEG | Intracranial stereotaxic EEG recordings |
| SL | Statistical learning |
| SMA | Supplementary motor area |
| SMG | Supramarginal gyrus |
| STG | Superior temporal gyrus |
| STM | Short-term memory |
| TP | Transitional probabilities |
| UF | Uncinate Fasciculus |
| WLI | Word learning index |
| wPLI | Weighted phase lag index |

Abstract

Language is one of the main traits of humans. Language is used to transmit information about thoughts, knowledge, and feelings in a very flexible way. Since childhood, humans show an innate ability to acquire language and to communicate. However, the easiness that children show to acquire language highly contrasts with the difficulty that occurs during adulthood. How both children and adults learn and develop language skills has been a matter of interest in many investigations. One of the milestones of language learning is vocabulary acquisition. This process requires the initial isolation of the novel word from the fluent speech and the subsequent association to a meaningful representation.

Research carried out in the last decades to study the mechanisms involved in language learning in adults and children provided wide knowledge of the process. Specifically, the development of neuroimaging techniques such as electroencephalography (EEG), magnetoencephalography (MEG), and magnetic resonance imaging (MRI) has allowed shedding light on the neural mechanisms that subserve word learning. Nonetheless, despite the advances in the field, there are still many unresolved questions about how the different learning steps co-occur on time and which brain regions are involved in each process.

The main goal of this Ph.D. thesis is to extend the current knowledge on the neural bases of language learning and vocabulary acquisition in both adults and children by using behavioral and electrophysiological measures. Specifically, we aim to understand better how the initial encoding and the gradual consolidation of novel word forms occur. We also aim to understand how speech segmentation and meaning mapping co-occur and how multimodal information is used in this process to facilitate the learning. Finally, we also want to directly assess the contribution of the hippocampus in the speech segmentation process. Here, we developed three experiments involving both children and adult populations and different experimental tasks to answer these questions and extend the current knowledge on the field.

Study I explored how novel word forms are gradually learned and integrated onto the mental lexicon. We developed two event-related potential word learning experiments in which novel word forms are associated with novel objects. In the first experiment, a group of 25 adults was exposed to the novel word-picture pairs over 5-day training sessions, while EEG was recorded during the first and last training session. This experiment allowed us to follow the neural traces

of the gradual build-up of mental representation for the novel items, from the initial encoding to the later consolidation of the word forms. This was observable by a decrease of the N400 amplitude within and between sessions and a shift of the topographical distribution from a frontal distribution the first day of training to a parietal distribution the fifth day. In *Experiment 2*, we aimed to replicate the first experiment and exclude possible confounds of the role of repetition and exposure in the task. With this aim, we randomly assigned 38 new participants to a consistent (Cons) or inconsistent (Incons) learning group. While in the Cons group, word-picture associations were consistent throughout the task; it was not the case for the Incons group. We showed that N400 modulations occurred only in the consistent group evidencing that results of *Experiment 1* derived from learning rather than a repetition effect. Overall, we showed that both fast and slow neuroplastic changes over frontal and temporal regions support the gradual build-up of new memory traces during online word learning.

In **Study II**, we explored speech segmentation and meaning mapping, two processes that have been separately studied in the literature. We also aimed to examine the role of multi-sensory cueing in both processes. With these aims, sixty 9-year-old children performed an audio-visual speech segmentation task while EEG was recorded. We showed that children, as adults, can simultaneously segment the words and associate them with meaning. Furthermore, our results showed increased neuronal activity at word and syllable frequency and enhanced fronto-occipital connectivity. The results suggest that multisensory cueing promotes that participants direct the attention to visual cues facilitating the word learning process.

In **study III**, we aimed to directly test whether the hippocampus is involved in speech segmentation based on statistical learning. We observed pharmaco-resistant epileptic patients' oscillatory activity with implanted depth electrodes while they performed a speech segmentation task. We showed that while auditory regions highly responded to syllable frequency, the hippocampus responded only to words. The finding provides direct evidence of the hippocampus's involvement in speech segmentation and advocates for a hierarchical organization of auditory information during speech processing.

Altogether, the thesis provides novel insights about how new words are extracted and encoded into the lexico-semantic network. The findings showed here may provide new insights to develop novel lines of investigation that improve our understanding of the neural and cognitive mechanisms involved in language learning and vocabulary acquisition.

Resumen

El lenguaje es una de las características únicas de la humanidad. Los seres humanos utilizan el lenguaje para transmitir información: pensamientos, conocimiento y emociones de una forma muy flexible e inigualable a ningún otro tipo de lenguaje. Sin embargo, la facilidad con la que los niños aprenden el lenguaje contrasta con la dificultad de esta misma tarea como adultos. Entender cómo los adultos y los niños aprenden nuevas lenguas ha sido tema de interés de muchas investigaciones en el campo tanto de la neurociencia como de la psicología. Uno de los puntos de inflexión en el aprendizaje del lenguaje tanto en adultos como en niños es la adquisición de vocabulario. Este complejo proceso requiere el aislamiento de la nueva palabra de entre el discurso del habla y su consiguiente asociación a un contenido semántico.

Investigaciones realizadas durante las últimas décadas han explorado los mecanismos neuropsicológicos implicados en el aprendizaje del lenguaje tanto en adultos como en niños. Específicamente, el desarrollo de técnicas de neuroimagen, como la electroencefalografía (EEG), la magnetoencefalografía (MEG) y la resonancia magnética (MRI), han permitido ampliar el conocimiento de los mecanismos neurales y electrofisiológicos de la adquisición de vocabulario. Sin embargo, a pesar de los grandes avances en el campo, hay todavía un gran número de cuestiones sin resolver.

El objetivo principal de esta tesis es ampliar el conocimiento actual sobre las bases neurológicas del aprendizaje del lenguaje y la adquisición de vocabulario tanto en adultos como en niños mediante técnicas de electrofisiología y paradigmas conductuales. Específicamente, el objetivo es entender mejor cómo ocurre la codificación inicial de las palabras y su consiguiente integración en el léxico. El segundo objetivo de esta tesis es explorar cómo la segmentación del discurso hablado y la asociación de nuevas palabras a un referente semántico ocurre, así como entender cómo la integración de información de forma multimodal influye en estos procesos para facilitar el aprendizaje. Finalmente, también tiene como objetivo explorar el papel del hipocampo en la segmentación del discurso del habla. Para ello, hemos desarrollado tres experimentos con diferentes tareas experimentales que hemos administrado tanto a adultos como a niños.

En el **Estudio I**, mediante un diseño longitudinal, hemos explorado cómo las nuevas palabras se van aprendiendo e integrando gradualmente al léxico. Para ello, hemos desarrollado dos

experimentos en los que nuevas palabras eran presentadas junto con nuevos objetos. En el Experimento 1, se expuso a un grupo de 25 participantes a asociaciones de palabras-objetos a lo largo de 5 sesiones experimentales mientras registramos la actividad eléctrica del cerebro el primer y último día. Este experimento nos permitió explorar los marcadores neurales de la formación de nuevas representaciones léxico-semánticas: desde la codificación inicial hasta la consolidación de las nuevas palabras. Esto fue observable por una disminución de la amplitud de la N400 tanto dentro de una misma sesión como entre sesiones. También observamos un cambio de la distribución topográfica de la N400, inicialmente frontal, pero que pasó a ser parietal el último día de entrenamiento. En el Experimento 2, quisimos replicar el primer experimento y a la vez confirmar que los resultados que habíamos encontrado no eran fruto de la repetición y exposición, sino del aprendizaje. Con este objetivo, asignamos de forma aleatoria 38 nuevos participantes a dos grupos, un grupo de aprendizaje consistente (Cons) y un grupo de aprendizaje no-consistente (Incons). En el grupo Cons las asociaciones palabra-objeto eran consistentes a lo largo de la tarea, mientras que en el grupo Incons las asociaciones iban cambiando en cada exposición de forma que dificultaban/impedían el aprendizaje. Los resultados confirmaron que la modulación de la N400 era debido al aprendizaje y no a la simple exposición, ya que ésta solo se produjo en el grupo Cons, mientras que en el grupo Incons no variaba. En definitiva, el conjunto de experimentos demostró que los cambios plásticos del cerebro, tanto los rápidos como los lentos, son la base de la adquisición de nuevas palabras.

En el **Estudio II**, exploramos la interacción entre la extracción de nuevas palabras dentro del discurso del habla y la asociación de estas nuevas palabras a su significado, dos pasos que típicamente se han considerado de forma independiente en el aprendizaje del lenguaje. También teníamos como objetivo explorar cómo la integración multisensorial influye en este proceso. Para ello expusimos a 60 niños de nueve años de edad a una tarea de segmentación audiovisual mientras registrábamos la actividad eléctrica del cerebro (EEG). Los resultados mostraron que los niños, igual que los adultos, pueden segmentar las palabras y asociarlas a su referente semántico de forma simultánea. Además, los resultados mostraron un incremento de la actividad neuronal a la frecuencia de las palabras y de las sílabas, así como un aumento de la conectividad entre regiones frontooccipitales a esas frecuencias. Estos resultados sugerían que la integración multisensorial facilita que los niños dirijan la atención visual hacia las pistas visuales, facilitando así el aprendizaje de nuevas palabras.

El **Estudio III** tenía como objetivo evaluar el papel del hipocampo en la segmentación del discurso del habla mediante mecanismos de “aprendizaje estadístico”. Para ello registramos la

actividad eléctrica de pacientes epilépticos que tenían implantados electrodos intracorticales mientras realizaban una tarea de segmentación del habla. Los resultados mostraron que el córtex auditivo respondía claramente a las sílabas, mientras que el hipocampo respondía solo a las palabras. Este resultado demostró la implicación del hipocampo en la segmentación del discurso basada en aprendizaje estadístico. Estos resultados sugieren a su vez una organización jerárquica del procesamiento auditivo durante la segmentación del habla.

En conjunto, esta tesis amplía el conocimiento existente sobre cómo las nuevas palabras se extraen y se codifican en el léxico. Los resultados mostrados aquí proporcionan nuevos conocimientos e indicios para desarrollar nuevas líneas de investigación que mejorarán nuestro entendimiento sobre los mecanismos neurales y cognitivos implicados en el aprendizaje del lenguaje y en la adquisición de nuevo vocabulario.

Chapter 1

Introduction

Chapter 1 Introduction

1.1 General Introduction

On an autumn day in 1947, Cathy and Keith Hayes returned to their suburban American ranch house with their newborn infant girl, Viki (Hayes, 1951). After a few difficult days, Viki began feeding well and growing rapidly. She was a very quiet baby, sweet and affectionate, and loved to be held and tickled. She learned to walk early, entering a rambunctious phase and breaking numerous household objects, but eventually, her loving parents' gentle discipline bore fruit and she developed into a playful, obedient little girl. By the age of three, Viki could feed and bathe herself, eat with a spoon and drink through a straw, and help with cleaning. She was fond of looking at herself in the mirror and loved assembling jigsaw puzzles. She enjoyed playing on the backyard swing, climbing trees, and playing peekaboo with the neighborhood children. She was in many ways a normal young girl, with one major exception: Viki did not speak. Not a word. She was able to grunt, scream, and laugh, so her problem was not with vocalization in general; instead, it seemed to stem from a neural difficulty specific to spoken language. After consultation with experts, Cathy Hayes instituted a speech training regime, manipulating her young pupil's lips manually and rewarding her with treats whenever she approximated a word. Unfortunately, even these dedicated efforts were mostly in vain: Viki's "vocabulary" reached a plateau of three words (*mama*, *papa*, and *cup*), and even these attempts were breathy and inarticulate: poor imitations of normal English speech. Viki seemed tragically doomed to a life without speech. Fortunately, Viki's parents were not totally surprised or alarmed by her failure to achieve speech or language, because Viki was a chimpanzee.

(Fitch, 2010)

Language is one of the main traits that differentiates humans from other animal species, and its uniqueness remains in the lack of an equivalent communication system in the animal kingdom (Nettl, 2000). Other animal species have sophisticated communication systems of varying complexity across different modalities (including auditory, olfactory, or visual). However, those systems are not as developed as human language. Human language goes a step further in communication: we use concrete and abstract information concerning the past, present, and future, which allows for the expression of any thoughts. Human language represents and communicates meaning in a much more detailed, flexible, and complex way than any other communication system (Fitch 2010), thereby allowing language to play a central role in human behavior, and therefore, in societies. Human language uniqueness and its critical role for humans have been of great interest in history. Specifically, how did humans develop such a

Introduction

complex system that, at an early age, seems to have been learned naturally but paradoxically, in adulthood becomes a challenging task?

Despite the innateness of human language, we are born without the ability to speak. Interestingly, the term "infant" comes from the Latin word "*infans*" meaning "speechless". Infants take most of their entire first year of life, during which they are exposed to human language, to start producing their first words. Some novels such as "The Jungle Book" (Kipling, 1894) or "El jinete del silencio" (Gonzalo Giner, 2011) feature characters who grew up isolated from human society and were, by consequence, not adequately exposed to language during infancy, causing difficulties in acquiring language and fluent speaking. While these stories are purely fictional, there have been similar cases in real-life, such as the story of Kaspar Hauser or that of the so-called "wild children" (Malson, 1964). All of these examples, fictional or not, illustrate how, despite the innate human motivation to acquire language, individuals who are not properly exposed to language during infancy are unable to develop language spontaneously, or if they do, remain unable to reach the level of a typical native speaker. From this evidence, one might deduce that we should explicitly train children to learn language (Bloom, 1998). In most western societies, parents put great effort into teaching children new words by directing them the speech with the use of linguistic and non-linguistic cues to facilitate the learning – i.e., looking or pointing at an object-.

In many other societies, however, parents do not try to teach language to their children. Instead, infants learn through simple exposure. Thus, the question arises: is this explicit teaching actually necessary? It has been suggested that parents provide useful cues that can facilitate learning (Weisleder & Fernald, 2013). However, these cues may simply not be necessary (Bloom, 1998).

One of the turning points of language learning is vocabulary acquisition. This process requires isolating words from fluent speech and inferring their meaning, thus representing one of the two language acquisition milestones. Word learning occurs throughout the entire lifetime. However, the mechanisms used to learn the first words during infancy differ in several ways from those used during second language (L2) learning in adults. *In this work, we explore the behavioural and neural bases of how humans acquire novel words and how different steps of vocabulary acquisition interact with each other to drive vocabulary acquisition.*

While exploring word-learning mechanisms, we will occasionally encounter one aspect of language recurrently debated: To what extent language acquisition depends on domain-specific or domain-general mechanisms? Part of the interest in this issue derives from the fact that the

Introduction

answer has profound implications for understanding how language learning occurs. It can also help in the near future to develop better treatments for patients suffering from language impairments (Peñaloza et al., 2016; 2017; Coran et al., 2020).

Some authors claim that language development relies upon **domain-general** mechanisms and argue that language development relies on general cognitive processes applicable to different domains and not specific to language. For instance, Piaget (1954) proposed three domain-general mechanisms (assimilation, accommodation, and equilibration) that may underlie the development of several cognitive domains. In a similar vein, Tomasello (1986; 2000) suggested that children's linguistic abilities rely on general cognitive systems. Indeed, syntactic development depends on children's ability to understand partitions of scenes into events (Tomasello, 1986). In recent studies, it has been shown that cognitive control processes that rely on domain-general mechanisms are involved in language production (e.g., Roelofs & Piai, 2011; Strijkers et al., 2011) and language comprehension (see Palermo & Rhodes, 2007, for a review).

Other authors propose that language development depends on **domain-specific** processes instead. This view reflects the idea that language relies on mechanisms/brain structures specialized in handling linguistic information independently of other cognitive systems. These theories claim that when the system is damaged, children show language impairment (Tager-Flusberg, Boshart, & Baron-Cohen, 1998; Duchaine & Nakayama, 2006). An example of such a theory has been proposed by Markman (1990), who claims that language learning proceeds with the help of several constraints that allow the acquisition. In the same direction, Chomsky (1986) proposed that innate grammatical knowledge is required for proper language learning. Those theories are domain-specific in the sense that they apply only to linguistic information. Markson & Bloom (1997) tried to disentangle the issue by performing an experiment in which they asked both children and adults to learn (i) novel words and (ii) facts about the novel words. They found that both adults and children were equally good at learning the words and learning the facts about the objects. As the former task does not need linguistic skills, the authors claimed that language learning mechanisms might be domain-general. These results argued in favor of a combination of both domain-general and domain-specific mechanisms for language acquisition (Hollich, Hirsh-Pasek & Golinkoff, 2000). The following section describes the main theoretical models of word learning.

Introduction

1.1.1 Neural bases of language

Our understanding of how distinct brain regions are involved in language abilities initiated in the 1960s when researchers began to investigate patients with language disorders. In 1861, a French physician Dr. Paul Broca tracked the evolution of a patient with speech loss resulting from syphilis and made a detailed description of his symptoms. When the patient died, Broca performed an autopsy and discovered the patient had a lesion in the third convolution of the left frontal lobe. Based on this observation, Dr. Broca determined that this region was responsible for speech production. Following this first patient, he studied other patients with language impairments and found lesions encompassing many of the same brain areas. A few years later, a German neurologist Dr. Carl Wernicke named this brain region "Broca's area". Dr. Wernicke also found that not all language impairments were due to damage in Broca's area. He realized that lesions in the left posterior superior temporal gyrus (STG), a region now referred to as "Wernicke's area", resulted in language comprehension impairments.

Interestingly, Dr. Wernicke postulated the existence of a direct connection between these two brain areas (Wernicke, 1874) and proposed that a lesion at any point along these connections would cause language impairment (see **Figure 1A**). The development of neuroimaging tools such as EEG, MEG, structural and functional Magnetic Resonance Imaging (fMRI) has allowed researchers to reveal more details about the cortical organization of the brain networks that support human language (Price, 2012). This led to the emergence of new models of language processing that partially support and complement the first models proposed by Broca and Wernicke.

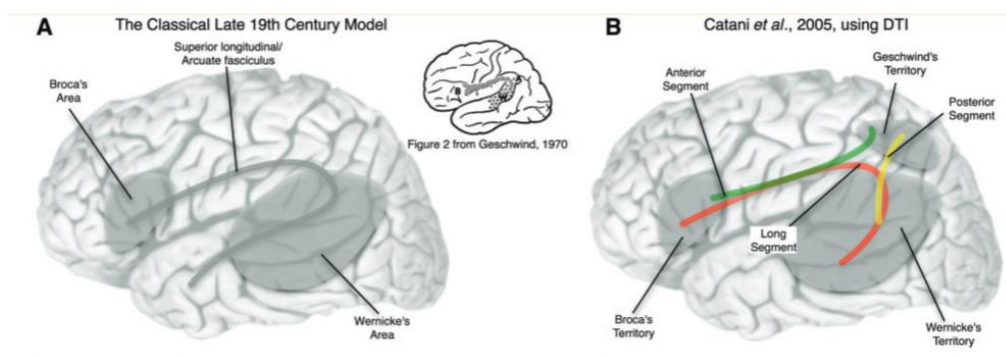
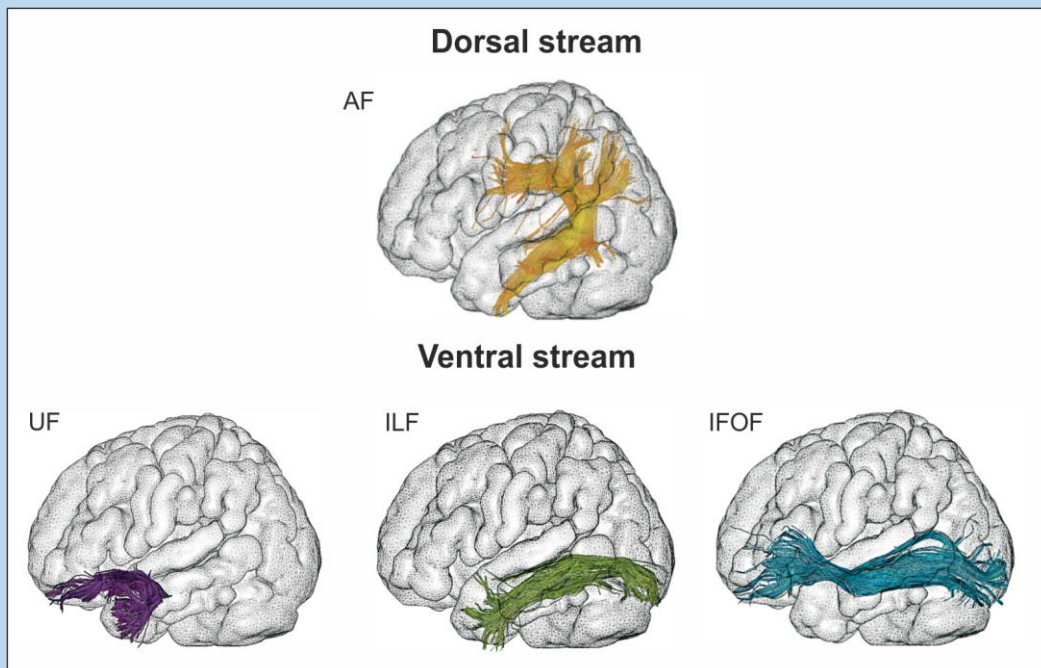


Figure 1. A. The classical view of brain connections of frontal and parietal language networks. B. Dorsal pathway. Model of AF organization proposed by Catani and colleagues (2005) (adapted from Dick & Tremblay 2012).

The dual stream model

The dual-stream model proposes two different routes for language processing: the dorsal and the ventral streams. The dorsal stream has been identified with the arcuate fasciculus (AF). The AF connects frontal (Broca's area), parietal (Geschwind's territory) and temporal (Wernicke's area) regions: the Broca's area, connects with Geschwind's territory through the *anterior segment*; Broca's area also connects with Wernicke's area, through the *long segment*; and Geschwind's territory connects with Wernicke's area through the *posterior segment* (Catani et al., 2005). The AF is considered as the main pathway for language processing. Specifically, it has been associated with word learning (François et al., 2016; López-Barroso et al., 2013) and verbal recall (Catani et al., 2007).

While the dorsal stream is composed of one unique pathway, several segments compose the ventral stream (see figure below). These ventral stream connect temporal regions with frontal areas. **The uncinate fasciculus (UF)** is a short connection between anterior temporal, and inferior frontal areas through the amygdala. It has been related to semantic processing (Gorno-Tempini et al., 2004; Papagno et al., 2011), and auditory working memory (Frey et al., 2004; Fritz et al., 2005; Diehl et al., 2008). It may also play a role in sound recognition (Clarke et al., 2002). The **inferior fronto-occipital fasciculus (IFOF)** connects the ventral occipital lobe and the orbitofrontal cortex. The **inferior longitudinal fasciculus (ILF)** runs in parallel, connecting nearby regions (Catani & de Schotten, 2008). Both fibers have been related to object recognition, face recognition, and visual-semantic memory (Milner & Goodale, 2008; Schmahmann et al., 2008; Turken & Dronkers, 2011).



Dorsal and ventral stream white-matter pathways (adapted from Catani and Thiebaut de Schotten, 2008). Dorsal stream: Arcuate fasciculus (AF, in yellow). Ventral stream: Uncinate Fasciculus (UF, in purple), Inferior longitudinal fasciculus (ILF, in green) and the inferior fronto-occipital fasciculus (IFOF, in blue).

Introduction

Initially, the scientific interest focused on the cortical regions involved in language (Weniche, 1874; Broca, 1861). More recently, however, the interest has shifted to the **connections** between these language areas. Indeed, brain connectivity patterns provide valuable information about how brain language areas may operate. The most outstanding neurobiological model of language processing is the **dual-stream model**, which proposes a left-lateralized dual architecture connecting frontal regions with parietal and temporal areas (Catani et al., 2005; see **Figure 1B**). The dorsal stream (indirect pathway) projects posteriorly through parietal regions up to temporal structures, while the ventral pathway directly projects through temporal regions. This stream may contribute to the mapping of auditory speech sounds to their articulatory representations, while the ventral stream may map auditory speech sounds to meaning (Hickok & Poeppel, 2000; 2004; 2007; Poeppel & Hickok, 2004; Rauschecker & Scott, 2009; Rodríguez-Fornells et al., 2009; see the box “The dual-stream model” for a detailed description of the model).

Although the dual-stream model builds a network upon the classical language brain regions, it only accounts for cortical structures (see Tremblay & Dick, 2016, for a discussion). This is a clear limitation as several fMRI and MEG studies suggest that different aspects of language processing may involve other structures like the medial temporal lobe (MTL, McNealy Mazziotta, & Dapretto, 2006, Henin et al., 2020), the basal ganglia (Price, 2012), or the cerebellum (Keren-Happuch et al., 2014; Pliatsikas et al., 2014).

1.1.2 Neural bases of language learning

In 2009, Rodríguez-Fornells and colleagues proposed an integrative model that delineated the major neural correlates of language acquisition during both first and L2 learning. Interestingly, this model provided an overview of the time course of word acquisition based on both electrophysiological and fMRI data obtained in various word learning paradigms. The model introduced a complementary view to the dual-stream model with the involvement of the MTL in language learning. Specifically, the model describes i) a **dorsal audio-motor interface** responsible for mapping sounds onto articulatory-based representations and phonological storage (Hickok & Poeppel, 2000; 2004; 2007, Wise et al., 2001). This interface includes the posterior STG, premotor regions, and Broca’s area involved in motor speech production. Notably, more recent work confirmed the involvement of the dorsal stream in the acquisition of new vocabulary in both adults (Hickok & Poeppel, 2007; Lopez-Barroso et al., 2013) and children (François et al., 2016). We will see that fMRI studies on speech segmentation also support this

Introduction

view (McNealy Mazziotta, & Dapretto, 2006; Cunillera et al., 2009; see section 1.2.1.1. The model also proposes ii): a ***ventral stream of meaning integration interface*** involved in the extraction of contextual information while using internal representations and previously acquired knowledge to infer word meanings. This stream seems to exhibit a more extensive scope. It involves areas in the superior and middle portions of the temporal lobe, including the inferior temporal gyrus (ITG), MTG, and inferior frontal gyrus (IFG, Hickok and Poeppel, 2007; Lau et al., 2008). MTG is proposed to reflect the activation of internal concept representations necessary to infer the meaning from the context (Hickok & Poeppel, 2007; Mestres-Missé et al., 2008; see section 1.2.2), while IFG seems to be involved in controlling the retrieval of semantic information (Badre & Wagner, 2002; Thompson-Schill et al., 1997). The stream has also been related to the creation of novel conceptual structures (Saur et al., 2008).

Importantly for the present work, Rodríguez-Fornells and colleagues proposed a third iii) ***episodic-lexical interface*** that involved the MTL, including the hippocampus, parahippocampal regions, entorhinal, and perirhinal cortices. They suggested that these regions may be responsible for the initial binding of new words into memory traces. Hippocampal regions have typically been associated with declarative memory, the storage of episodic memories (McClelland et al., 1995; Ullman, 2001 see section 1.2.2.2), and the extraction of patterns regularities from the environment (Turk-Browne et al., 2009; see section 1.2.1.2). Thus, it seems plausible that hippocampal regions are also involved in the extraction of language regularities and the storage of novel information into memory. Finally, they also uphold that other structures such as the basal ganglia may modulate the information flow between the cortex and thalamus structures, which, in turn, may modulate the interaction between the three interfaces (Haber & Calzavara, 2009).

1.2 Language learning mechanisms

1.2.1 Word segmentation and statistical learning

The first language (L1) ability that infants seem to develop between 1 and 6-months-old is the capacity to perceive phonemes as organized into a system of categories (Eimas et al., 1971). An example of this, in Catalan, are the words *os* (which means bear) and *ós* (which means bone). These two words can be distinguished in written form using the diacritical accent and orally through the different pronunciations of the letter [o]: /o' / and /ɔ'/. This phonological contrast is easily differentiated by a native Catalan speaker but not by a Spanish speaker due to Spanish

Introduction

only having the phoneme /o/ for the letter [o]. Phonemes are language-specific, and thus, while infants between 6 and 10 months of age have little difficulty in distinguishing phonemes regardless of the language (Aslin et al., 1981), adults that do not have the specific phonetic contrasts in their native language have difficulties in discriminating these sounds (Werker & Tees, 1984; Lisker & Abramson, 1971).

Phonemes appear structured in the speech signal, forming syllables. It then follows words that are made up of one or more syllables. Identifying these words in fluent speech is crucial for language learning for both adults and infants, as words are the units that carry meaning in language. While word segmentation is easily accomplished in written texts where spaces separate words, it becomes more difficult in fluent speech, where word boundaries are not always denoted by silences (see **Figure 2**; Liberman et al., 1967). The **speech segmentation problem** is indeed one of the first problems that language learners must solve when learning a new language. Interestingly, infants seem to be more proficient in this task than adults, with decreasing performance after puberty (Johnson & Newport, 1989; Weber-Fox & Neville, 2001).



Figure 2. Speech waveform from the sentence “There are no silences between words” that illustrates the lack of boundaries between words in normal speech (from Saffran, 2003)

To solve this problem, learners tend to use linguistic and non-linguistic cues to extract word units from continuous speech (Kuhl, 2004; Davis & Gaskell, 2009; Rodríguez-Fornells et al., 2009). Across all languages, speech sounds appear in combinations following general rules: the phonotactics (for sounds) and orthotics (sound-letter equivalence) of a language. **Phonotactics** refer to the restrictions on combinations of phonemes within and between syllables (i.e., consonant-vowel combinations or consonant clusters). Phonotactic rules can be used to segment words (Mattys & Jusczyk, 2001). Similarly, **prosodic and rhythmic cues** are used to identify words and sentence boundaries. For instance, **stress patterns** that refer to a specific syllable's emphasis within a word are also important for speech segmentation (Sanders & Neville, 2000; Cutler & Carter, 1987; Jusczyk, Houston, & Newsome, 1999). Stress patterns are

Introduction

common in many languages. However, some languages, such as Finnish or Dutch, have fixed initial stress cues, which provides a more useful cue to segment words than in other languages such as Spanish that have a variable stress pattern (Toro-Soto, Rodríguez-Fornells, & Sebastián-Gallés, 2007). In addition to linguistic cues, learners can also use other cues such as prior knowledge or contextual information to solve the problem. There are, overall, several cues that learners can use. Nonetheless, preferences for different cues seem to change across the lifespan (i.e., infants strongly rely on acoustic cues over distributional information, Johnson & Jusczyk, 2001; Thiessen & Saffran, 2003; Hollich, Hirsh-Pasek & Golinkoff, 2000, see section 1.2.2).

The learners can also extract regularities from the environment to acquire knowledge, and particularly language. Among empiricist researchers, Reber (1967) is known for his work on **implicit learning**. Implicit learning is defined as the process by which knowledge about complex stimulus domains is acquired largely independent of conscious awareness of either the process or the products of acquisition (from Reber & Allen, 2000). Reber conducted an experiment in which participants were exposed to strings of letters forming a complex artificial grammar and showed that participants were sensitive to the underlying artificial grammar. Furthermore, participants were able to use the acquired grammar information in a recognition task. Reber highlights the participants' innate sensitivity to stimuli without the need for explicit learning strategies, emphasizing its close relationship with the concept of **perceptual learning** proposed by Gibson and Gibson (1955). Perceptual learning refers to how an organism perceives and responds to the environment in a reliable and efficient manner (Gibson & Gibson, 1955). They argued that both concepts refer to an intrinsic and inductive process in language learning.

Saffran and colleagues (1996) were the first to propose that learners use **transitional probabilities** (TP) to segment words. TP refers to the probability that one sound follows the next is higher within words than between words (Swingley, 2005). For example, in the sentence '*the computer is there*', the probability of the string '*com*' be immediately followed by '*puter*' is the same as the probability that the string '*ter*' be followed by '*is*'. However, if we use the same word in another sentence such as '*The computer does not work*' the syllable '*com*' is still followed by the syllables '*puter*' but the syllable '*ter*' is followed by '*does*' instead of '*is*'. TP can be computed as

$$TP = P(Y/X) = \text{Frequency}(XY) / \text{Frequency}(X)$$

the probability of Y appearing when X is present is calculated by dividing the number of times Y appears with X (Frequency (XY)) by the total number of times X has appeared (frequency (X)). In

Introduction

the example above, the TP of the syllable 'com' would be computed by dividing the number of times that the syllable 'com' is followed by the syllable 'puter' by the total number of times the syllable 'com' appears. As the syllable 'com' is followed by 'puter' in both sentences, the TP of 'com' being followed by 'puter' is 1.0. If we instead compute the TP of 'ter' being followed by 'is', the TP would be 0.5.

In their landmark study, Saffran and colleagues (1996) demonstrated that infants computed TPs between syllables and used distributional information from the speech stream to segment words. They created 2-minutes of continued synthetic speech stream composed of four tri-syllabic pseudo-words (e.g., badaku, padoti) presented an equal number of times in random order but with equal probability of appearance regardless of the previous word. There were no acoustic cues at word boundaries, and thus, only TPs across words could be used to segment the stream. As the words were tri-syllabic, the TPs between syllable pairs was higher within words (1.0) than between words (0.33). In their experiment, 6-month-old infants were exposed to the stream and afterward tested on their ability to recognize stream words when the words were presented together with novel tri-syllabic words containing the same syllables as the stream words but not originally presented in the stream. The results showed that infants were able to recognize the stream words, and thus that they used the statistical regularities in the speech stream to segment words from the stream. Infants' ability to use statistical information in the environment to bootstrap language acquisition was termed **statistical learning (SL)**. This process is proposed to occur incidentally, without awareness, and as a product of simple exposure (Saffran et al., 1997; Turk-Browne et al., 2005; Saffran et al., 1999).

After Saffran's experiment, SL became a matter of great interest. The initial experiment has been replicated (Aslin et al., 1998) in different ages, sensory modalities, and even species. These studies have demonstrated that SL still occurs even when the task is complicated by factors such as added conflict stress cues (Johnson & Jusczyk, 2001), and when using either naturally produced syllables (Graf-Estes et al., 2007; Sahni Seidenberg & Saffran, 2010), or complex natural language (Pelucchi, Hay & Saffran, 2009a; 2009b, Hay et al., 2011). These studies have also found that infants can track statistical regularities when using different word lengths (Johnson & Tyler, 2010) or non-adjacent dependencies (Newport & Aslin, 2004). Further studies showed that SL is maintained throughout adulthood (Saffran et al., 1999; see next section). It has also been shown that it is present in some animals in a primitive form (Toro-Soto & Trobalón, 2005; Meyer & Olson, 2011). Moreover, SL is not restricted to language, and both infants and

Introduction

adults can track non-linguistic auditory (Saffran et al., 1999) and visual statistical regularities (Fiser & Aslin 2002a; 2002b).

Infants continue to develop their segmentation abilities during the first years of life while at the same time they begin to associate words with meanings, exploiting the statistical structure of speech to succeed at word learning (Graf-Estes et al., 2007; Hay et al., 2011). Interestingly, previous research on early word segmentation has shown that both electrophysiological and behavioral data during infant speech segmentation task predicted later language outcomes during the development. Specifically, it has been shown that word segmentation before 12-month-old predicts both expressive and receptive vocabulary of one-year-old and above (Newman et al., 2006; Junge et al., 2012; Singh et al., 2012; François et al., 2016). These studies suggest that developing good SL abilities during infancy is determinant to proper language development in later stages of language learning.

1.2.1.1 Neuroimaging studies on speech segmentation based on SL

The introduction of the fMRI technique in speech segmentation studies shed some light on the brain regions involved in the process. McNealy, Mazziotta, and Dapretto (2006) were the first to use this approach. Following the procedure used in Saffran and colleagues (1996), they constructed three nonsense speech streams: one with stress cues, another without stress cues, and the last one composed of random syllables. They analyzed participants' brain responses to passive listening to the streams and found an overall activation of language-related areas, specifically, bilateral STG, left supramarginal gyrus (SMG), and some regions of the frontal cortex (see **Figure 3**). They also found a reduction in brain activity throughout the task, which they interpreted as increasing efficiency. They suggested that the STG activation reflected the ongoing computation of statistical regularities and that SMG activation marked the development of phonological representations instead. Although these modulations of brain activity were observed during the exposure to the streams, participants could not consciously detect the words in a behavioural test, highlighting the implicitness of the task. Interestingly, the same researchers performed a study using a similar paradigm with typically developing children and children with autism. While in normal children, the patterns of activation were similar to those observed in adults, children with autism showed a reduced activity (see **Figure 3**). The hallmark feature of autism disorder is a delayed language development (Bailey, Philips & Rutter, 1996); thus, the results suggested that those regions were directly related to normal segmentation abilities and language-related learning.

Introduction

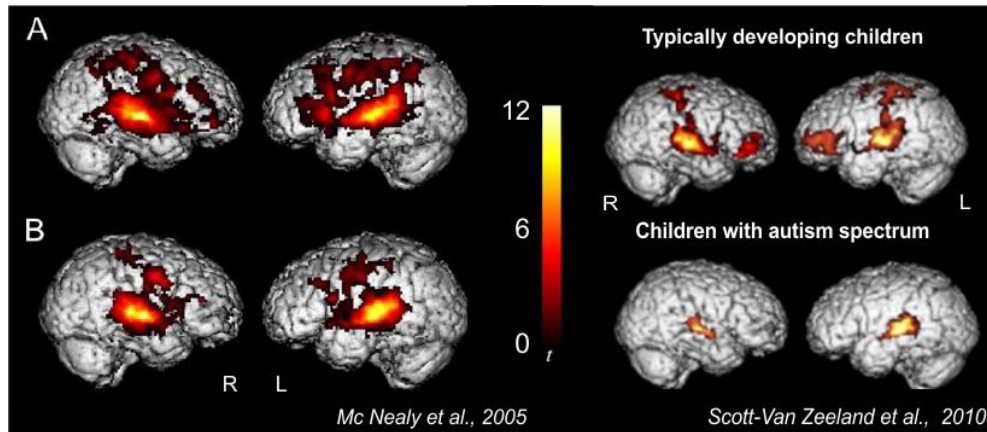


Figure 3. The pattern of activation found in McNealy Mazziotta, & Dapretto, 2006 for the unstressed (A) and stressed language stream (B) and for typically developing children and children with autism for the unstressed condition in Scott-Van Zeeland et al., 2010.

Afterward, Cunillera and colleagues (2009) went one step further by combining EEG and fMRI techniques to explore the time-course of speech segmentation. Participants were exposed to four one-minute random streams and to four one-minute statistically structured stream. They found activations in the posterior STG and ventral premotor cortex (PMC) during the first two minutes of learning (referred to as the first block). Interestingly, superior ventral PMC activity was no longer significant in the second block (later two minutes) of exposure (**Figure 4**). They proposed that this activation reflects the role of this region in mapping the acoustic properties of phonemes its involvement in the process of segmentation. Moreover, these results provided further evidence for the implication of the PMC in speech perception (Fadiga et al., 2002; Wilson et al., 2004).

Following this work, Karuza and colleagues (2013) conducted an experiment in which exposed participants to two forward and two backward streams (that consisted of playing the forward stream in reverse). They compared brain response to both types of streams with a third type of stream, an overlaid stream, that consisted of layering three copies of the forward stream. They reported stronger activations of the left STG, as shown by McNealy and Cunillera's experiments; however, they also found activations in the left IFG (**Figure 5**). This brain region has been related to domain-general SL (e.g., in Turk-Browne et al., 2009, for a visual paradigm) and thus has been proposed to be modality-independent. Interestingly, simultaneous activation of IFG and STG has also been found in Abla and Okanoya's experiment (2008), which carried out a similar paradigm but with tones rather than linguistic stimuli (with near-infrared spectroscopy (NIRS) technique).

Introduction

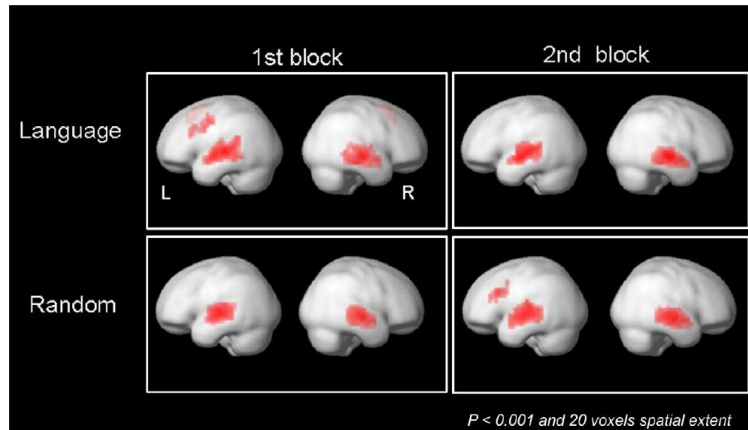


Figure 4. Pattern of functional activations found in Cunillera et al., 2009 for the first and second blocks in both language and random streams.

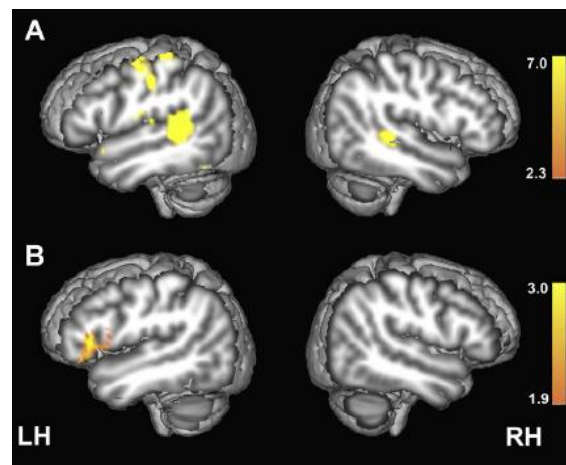


Figure 5. The pattern of results found in Karuza et al., 2013. showing the functional activations for the contrast forward > backward stream in the whole-brain analysis (A) and the specific neural changes in the IFG (B).

In another important study, López-Barroso and colleagues (2013) provided multi-methodological evidence for the involvement of the AF in speech segmentation based on SL. They used diffusion tensor imaging (DTI) and fMRI while participants performed a speech segmentation task on a stream composed of tri-syllabic words separated by brief pauses. They found that the fronto-temporal connectivity strength, that is, the functional connectivity between language regions (Broca's and Wernicke's area), predicted the level of learning. Furthermore, they showed a positive relationship between word learning performance and microstructural properties of the AF. In a subsequent experiment, López-Barroso and colleagues (2015) used a data-driven approach based on independent component analysis (ICA) to identify the functional brain networks that subserve word learning (**Figure 6**). They showed that the frontoparietal network, the auditory-premotor network, the sensory-motor network, and the

Introduction

frontotemporal network were consistently activated during a word learning task from an artificial language speech stream. These results provided further evidence for the importance of frontoparietal connectivity patterns in speech segmentation based on SL as well as the participation of other domain-general neural networks, most probably related to attention and cognitive control (Sliwiska et al., 2017).

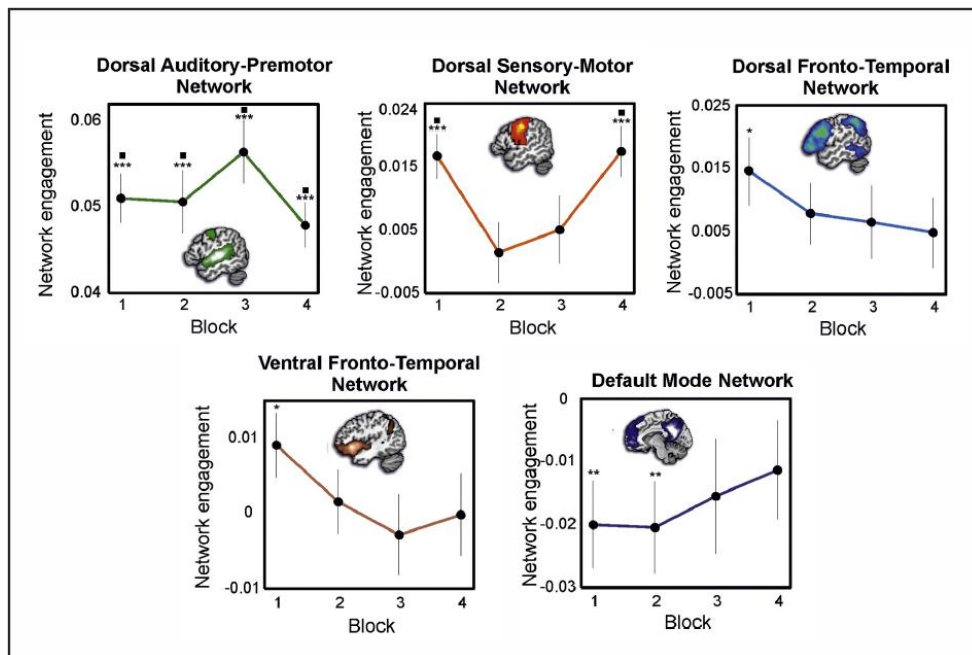


Figure 6. Results from Lopez-Barroso et al., 2015 showing the engagement of the functional networks subserving word learning along with the block): the dorsal auditory-premotor network (green), the dorsal sensory-motor network (red), the dorsal fronto-temporal network (blue), the ventral fronto-temporal network (yellow) and the default mode network (dark blue).

To summarize, the neuroimaging studies described above have provided converging evidence for the involvement of specific cortical brain areas during auditory and visual SL. Overall, the studies showed that although bilateral activations were found, there was also great left-lateralization involving the temporal regions such as STG (McNealy Mazziotta, & Dapretto. 2006, Cunillera et al., 2009; Karuza et al., 2013), frontal regions such as the IFG (Karuza et al., 2013) and PMC (Cunillera et al., 2009). These regions correspond to the dorsal route first proposed by Hickok & Poeppel (2007) in their dual-stream model of speech processing and subsequently included in the language learning model from Rodríguez-Fornells and collaborators (2009).

1.2.1.2 SL and the hippocampus

The neuroimaging studies mentioned above proposed a widespread number of brain structures that seem to be involved in SL. These structures include modality-specific cortical networks such

Introduction

as auditory regions (McNealy Mazziotta, & Dapretto, 2006; Cunillera et al., 2006) but also structures involved in domain-general cognitive processes such as basal ganglia (Karuza et al., 2013). Nonetheless, one of these studies' shared conclusions is a clear relation of MTL structures with SL. Among these structures, the hippocampus has been taken a special interest. Turk-Browne and colleagues (2009) were the first to suggest the involvement of the hippocampus in SL. In their study, participants were visually exposed to statistically structured or unstructured sequences of shapes while performing a task unrelated to the structure in an fMRI session. Neural responses to MTL, as well as basal ganglia (striatum), were found when participants were exposed to the structured but not to the unstructured task. These findings seem to fit well with the declarative/procedural model proposed by Ullman (DP; 2001; see section 1.2.2.2.3).

In a single-case study, Schapiro and colleagues (2014) showed that a patient with MTL and bilateral hippocampal loss (**Figure 7**) was unable to detect the statistical regularities when exposed to a stream in various modalities (continuous sequence of shapes, syllables, scenes, and tones). Although the study suggested that hippocampal damage completely inhibits the capacity of SL, a replication of the study with a larger sample size of patients (Covington, Brown-Schmidt & Duff, 2018) showed that although the hippocampus is necessary for SL, hippocampal lesion alone does not abolish SL. In another set of experiments, individual differences in hippocampal volumes were related to performance on visual SL in adults (Schlichting et al., 2017). The results suggested that hippocampal volume differences were related to SL ability but also with general associative tasks in adults (Schlichting et al., 2017) and expressive language abilities in children (Lee et al., 2015). In the same line as the former study, Piai and colleagues (2016) provided evidence of the hippocampus's involvement in speech processing. They exposed participants to sentences with a constrained or unconstrained context and showed that hippocampal theta oscillations increased in constrained conditions compared to the unconstrained context.

Introduction

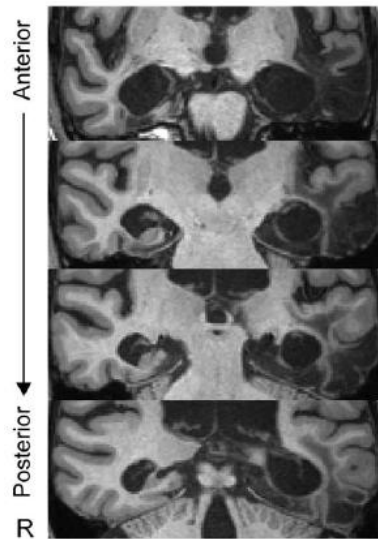


Figure 7. Display of the lesion location of the patient from Schapiro et al., 2014 experiment. Coronal slices from T1-weighted images showing a lesion in both hippocampus hemispheres, bilateral damage to other MTL regions, and damage in the left anterior temporal lobe.

Overall, this body of evidence suggests that the hippocampus is involved in the extraction of pattern regularities across different modalities. However, although the evidence showing that the hippocampus is involved in SL in the visual domain seems quite consistent (Turk-Browne et al., 2009; Schapiro et al., 2012), the involvement of the hippocampus in auditory SL is less clear. The evidence reported by Schapiro et al., 2014 comes from a single case study, and the posterior replication with a higher sample size showed that although it seems to be involved, maybe it is not completely necessary. Few studies have related hippocampal activity with language processing (Duff & Brown-Schmidt, 2012; Piai et al., 2016). Despite these few studies, the role of the hippocampus in language is not completely understood and deserves further consideration.

1.2.2 Word learning and vocabulary acquisition: the world-to-word mapping problem

Once words are extracted from fluent speech, learners face another problem: associating the novel word-forms with a meaning (integrating the novel word-form in the semantic network). The acquisition of meaning through vocabulary learning is the core of language learning (Bloom, 2000). When a child encounters a new word for the first time, he must memorize the word's phonology and, at the same time, infer its meaning. For this process, learners use linguistic and non-linguistic information (i.e., visual cues) to rapidly create new lexical-semantic associations (Gillette et al., 1999; Lindsay & Gaskell, 2010). Importantly, this process changes over time: the

Introduction

mechanisms used to learn the first words of the mother tongue during childhood differ in several ways from those used during L2 learning in adults (Ullman, 2001). Despite the differences, the literature on child language acquisition has been a useful tool for studying the meaning acquisition process.

Infants begin to say their first words between 8 and 12 months, suggesting that they are already able to make phonological-semantic associations by this age. During the second half of the child's second year, there is a complete shift in word learning rate from slow to a fast learning rate (Nelson, 1973; Markman, 1990; see **Figure 8** for an illustration). This switch called the *vocabulary spurt*, allows infants to learn the meaning of novel words very fast and with only a few exposures (Friedrich & Friederici, 2008). From that moment, children seem to be able to fast map a novel word into meaning, a phenomenon called *fast mapping* (Carey & Bartlett, 1978).

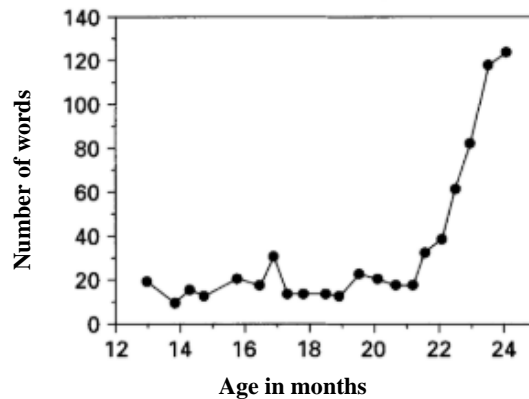


Figure 8. Example of the vocabulary spurt that occurs during children's second year of life (Barrett, 1999; adapted from Plunkett, 1993).

Fast mapping was first described in the seminal study by Carey & Bartlett (1978), in which children were presented with a blue and an olive tray. The experimenter asked the children, "bring me the *chromium* tray, not the blue one, the chromium one", and all the children correctly bring the olive tray, thus, inferring that chromium was the name of the new color (olive). Interestingly, most children remembered the name of the color when retested one week after the experiment, and some of them even after a month.

While fast mapping explains how children can assign meaning to new words using very few exposures, other authors have proposed that the meaning mapping problem can be solved using information from multiple encounters with the concept/word to be learned (Smith et al., 2014), that is, by merely domain-general **associative learning mechanisms**. This theory is based on the idea that when a word co-occurred with an event or a stimulus, an association is formed. Smith

Introduction

(2000) proposed that word learning is the consequence of an associative process in which new words are related to the most salient referent within the context. Yu and Smith showed that both adults and infants exposed to trials in which multiple objects co-occurred with new words learned the object labels by computing cross-trial statistics (Yu & Smith, 2007, 2011). Notwithstanding, this view cannot explain how word learning occurs in a difficult situation where the learner faces multiple new words and multiple possible referents.

Quine (1960) discussed the complexity of language learning using a thought experiment, **the “Gavagai” parable**: Imagine that an anthropologist is visiting a group of natives to study their language and behavior. One day, during his visit, a rabbit passes by, and one of the natives exclaims: “gavagai!”. The anthropologist then probably assumes that “gavagai” means rabbit. Nonetheless, Quine claims there is no way for the anthropologist to be sure about the word’s meaning. He argues that “gavagai” could refer to the tail of the rabbit, its skin color, or even animal type (herbivore, prey). Despite the problem raised by Quine, children generally correctly guess a word’s meaning with just one-word exposure along with the target object, where they seem to assume that the novel labels refer to the whole object (Carey & Bartlett, 1978). This preference for assigning the name to the entire object is not exclusive of language; it also applies to other domains such as counting. Shipley & Shepperson (1990) asked 3- and 4-year-old children to count the number of forks presented to them, among which there was a broken fork which only had half of the fork. Participants included the broken fork in the sum, thus showing the “whole object preference” in non-linguistic domains. A possible explanation of this phenomenon is that children and adults do not seem to process all the possible meanings of a word (or possibilities, in the case of counting). Still, instead, they seem to constrain its possible meanings by simplifying the problem.

Most recent theories agree on **pre-existing constraints** in word learning (Clark & MacWhinney, 1987; Markman, 1990). That is, the fact that *language-specific mechanisms bias the way we acquire language*. Those theories stand that innated ways of interpreting the information have come to determine how children acquire language (Barrett, 1990). Despite the great acceptance of this idea, the nature and number of those constraints remain under debate. Markman (1990), one of the prominent supporters of this idea, proposed that children acquire novel words constrained by three different factors based on how words link to categories: i) the whole object assumption (Carey & Bartlett, 1978), ii) the taxonomic assumption which leads children to consider words as referring to objects of like kind, rather than to objects that are semantically related and iii) the mutual exclusivity assumption which leads children to expect that objects

Introduction

have only one label. Markman and Wachtel (1988) tested their hypothesis by presenting 3-year-old children with novel objects with a salient part while the experimenter labeled the object. They found that if the whole object did not already have a name, children associated the label with the entire object (whole object bias). If the object already had a familiar name, children associated the meaning with the object part (mutual exclusivity assumption).

Another constraint that children (and learners in general) show during word learning is their tendency to extend the word to other objects of the same kind, the so-called **generalization problem**. In a behavioural experiment, Gelman and Markman (1986) explained to a group of children that a “brontosaurus” had cold blood while “rhinoceros” had warm blood. The children were then asked to guess the blood-type of a “triceratops”. As “triceratops” shows more similarities with rhinoceros than with brontosaurus, children mostly inferred that triceratops had warm blood. However, when children were told that “triceratops” is a dinosaur similar to the brontosaurus, children inferred that “triceratops” had cold blood. This experiment showed that children use the labels (in this case, dinosaurs) to attribute them to the same category. However, when they do not have categorical information available, they use perceptual properties to generalize the labels. Similarly, children also seem to generalize object labels for objects having the same function (Kemler Nelson et al., 2000), thus suggesting that general intuition and knowledge about objects are used to generalize the labels (Bloom, 2000).

Other theories on word learning have centered on the social context of communication as a basis for word learning. Vygotsky, a Russian psychologist (1980), proposed that although biological factors constitute a pre-requisite for language development, sociocultural factors are indispensable. In the same line, Tomasello and colleagues (1995) proposed that word learning is driven by social interactions. Children’s knowledge of people’s goals and intentions limits the possible interpretations of the meanings of novel words. Another theory based on social insights was proposed by Kuhl (2004), who suggested that children learn words by detecting statistical probabilities presented in a social context. Bloom (2000) proposed that children conceive the world as containing objects, properties, and events driven by the whole-object bias. He also suggested that only a few months after their first word, children have enough understanding of language to exploit the semantic and syntactic relationships (linguistic cues) to find out words meaning, just as an adult would do. He also argued that children know enough about others’ minds (i.e., the theory of mind) to figure out what they refer to when speaking. Under his perspective, constrained context and social interactions are the key factors in children’s vocabulary acquisition. These factors may direct children’s attention to make them focus on the

Introduction

correct referent. This idea is supported by experiments showing that children actively search for a speaker's gaze to identify the intended referent (Akhtar, Carpenter & Tomasello, 1996; Bloom, 2000; Tomasello & Barton, 1994) and that children do not attach the label to an object if the speaker does not seem to be referring to that object (Baldwin, 1995). Overall, social-pragmatic or *social-gated language learning theories* argue that children learning capacities *capitalize on the social environment in which normal children develop*.

Despite the existence of many word-learning theories in children, researchers increasingly agree on the necessity to consider multiple factors from different domains to explain the phenomenon. Hollich, Hirsh-Pasek & Golinkoff (2000) proposed a *compiled model positing that children's word learning ability emerges from the combination of associative learning mechanisms, word learning biases (constraints), and social-pragmatic factors*. They argue that children use both linguistic and non-linguistic cues for learning new words. Notably, the weight given to those cues may change through development. During the initial stages of word learning, children would rely more on domain-general cues (e.g., social and attentional cues). Conversely, during later learning stages, they would rely more on language-specific cues (e.g., phonotactic or grammatical rules). They proposed that children's word learning is the product of multiple factors with different weights through development (**Figure 9**).

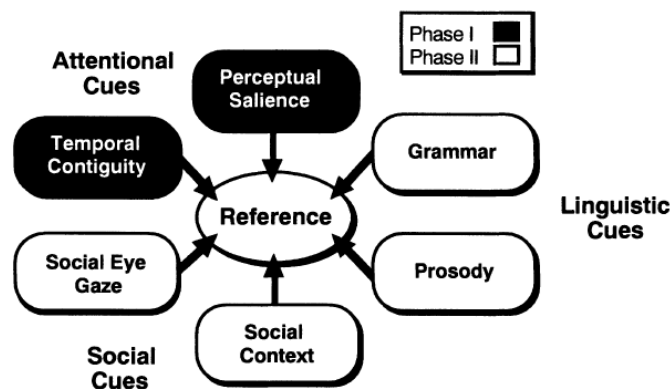


Figure 9. Model proposed by Hollich, Hirsh-Pasek & Golinkoff (2000) about a change in children's cue's reliance across development. Initially, children rely on attentional cues such as temporal contiguity or perceptual salience (Phase I). Cue's reliance later shifts towards linguistic and social cues (Phase II).

In the same direction, Räsänen and Rasilo (2015) proposed a combined model of word segmentation and meaning acquisition mediated by cross-situational learning. Based on empirical and simulated data, the authors proposed that initial domain-general mechanisms based on SL may facilitate vocabulary acquisition across modalities. Specifically, the learners

Introduction

would use the statistical cues found in the speech signal (e.g., TP between and within words) and in the external context (e.g., the co-occurrence of a word-form with its referent through multiple encounters) to isolate and provide meaning to the novel word-form. This theory is supported by experiments that showed that multi-modal cues facilitate both speech segmentation (Cunillera et al., 2010b; François et al., 2017b) and meaning acquisition (Smith, 2000; Plunkett, 1997). Overall, language learners (both adults and children) may use a variety of approaches (SL, associative learning) across different modalities (visual, auditory) to parse fluent speech and to map novel word-forms to their referent. Hence, these two language learning steps would not occur independently, but instead, they may influence each other to facilitate the learning process.

1.2.2.1 Lexico-semantic organization

To acquire vocabulary, children must have a mental representation of the relevant information about the word (concept). The mental lexicon consists of a set of all these representations that we have in mind. Understanding how these concepts are stored and organized in the brain is essential to understand how we integrate novel word-forms and their associated meaning during word learning. Collins and Loftus (1975) proposed that word meanings are represented by conceptual nodes connected by semantic relations or associative relations between them (see **Figure 10**). This model assumes that activation spreads from one concept to another and that when a node is activated, the closely connected ones are activated too. The consolidation of a new word into the mental lexicon might require binding this new label onto existing semantic information (Tamminen & Gaskell, 2013). Thus, the integration of new words should be better when the novel word shares many semantic features with the existing concepts (Perfetti, 2007). Importantly, this specific step may differ in adults compared to children. Differentially to adults that have a large lexico-semantic network, young infants' mental representations are under-developed. Therefore, the semantic relationships they build may be more elementary and need to be shaped with time. It is also worth reminding that L1 acquisition in infants develops parallel with other cognitive functions (Diamond, 2002). Therefore, some language-related cognitive mechanisms are also part of the natural learning and development of a child's cognitive abilities (i.e., memory and attentional systems). Understanding how these processes interact with language-learning mechanisms during development seems crucial to comprehend better the language acquisition process.

Introduction

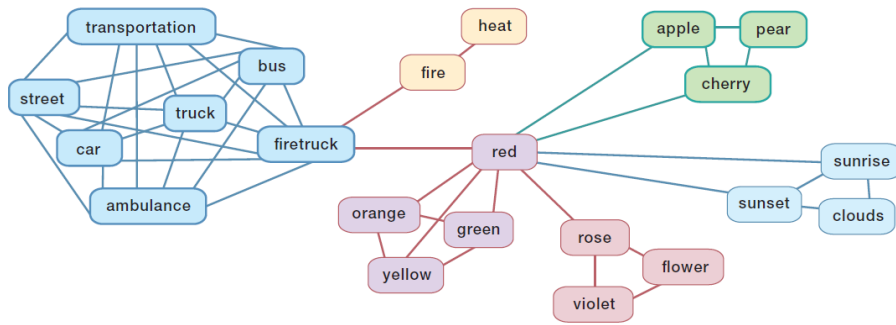


Figure 10. Representative map of semantic network organization (Gazzaniga, 2014). The figure illustrates the semantic network for the word “firetruck” and its lexico-semantic associations with other concepts. For example, firetruck is connected with red because firetrucks are generally reds.

Another important issue is how do we integrate novel word-forms into the mental lexicon during L2 learning. While L1 acquisition seems to occur effortlessly and implicitly, L2 learning is considered a non-automatic and high demanding task (Newport, 1990; Snow & Hoefnagel-Hohle, 1978; Rodríguez-Fornells et al., 2009). *One of the main differences between L1 and L2 is that in L1, novel words are associated with new meanings, but in L2, word-forms are associated with an already existing meaning.* An unresolved question is whether the links between novel L2 word-forms and concepts are created by direct connections with concepts or if they are formed through lexical links with L1. Kroll and Steward (1994) proposed that, in the early stages of L2 acquisition, novel words are attached to meaning via lexical links with L1 words. Still, as learners become more proficient in L2, direct links between L2 words and concepts are built (see **Figure 11**).

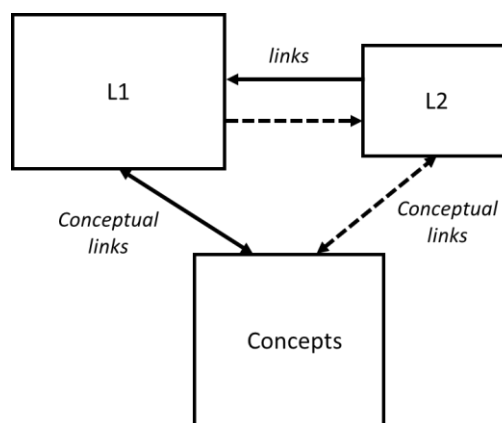


Figure 11. A revised hierarchical model of lexical and conceptual representation in the bilingual model (an adaptation from Kroll & Steward, 1994).

Introduction

1.2.2.2 The role of memory in language acquisition

During a conversation, a speaker produces roughly 150 words per minute (Studdert-Kennedy, 1987), causing words and sentences to rapidly superimpose and generate massive amounts of information over a short period. This amount of information makes listeners quickly lose track of the words they just heard and retain only a few pieces of information. Besides, our ability to maintain auditory sequences in memory is limited; overall, humans have limited capacity for recalling the temporal sequence of both auditory and visual information (Warren, Welch & McCarthy, 1981; Wilson & Emmorey, 2006). How does the brain store so much information? What causes some words to be easily retained while others are not? Despite the difficulty of this task, some information is kept in memory and can be retrieved after days or even years. Several models have been proposed to explain how memory and language interact.

1.2.2.2.1 The phonological loop

Baddeley and Hitch (1974) proposed a model, termed *the phonological loop*, explaining how short-term memory (STM) and working memory mediate the initial steps of vocabulary acquisition (**Figure 12**). For language processing, STM is needed to temporarily store the strings of sounds and words necessary to process a sentence (Baddeley, Papagno & Vallar, 1988). STM is also needed for the initial retention of word-forms and its meaning. *The phonological loop* contains two different components: i) a phonological store that holds verbal stimuli immediately into a phonological form but with a limited time-capacity, and ii) a rehearsal process (motor) that serves to refresh information and avoid decay (Gathercole & Baddeley, 1993) of the representations. Research on children supporting this proposal has shown that a better performance in digit span and non-word repetition test positively correlates with the ability to learn new words, as they both strongly rely on working memory capacities (Gathercole & Baddeley, 1989; 1990). Other studies also have shown that non-word repetition and phonological learning are good predictors of language learning (Service, 1992; Kohonen & Service, 1995).

Introduction

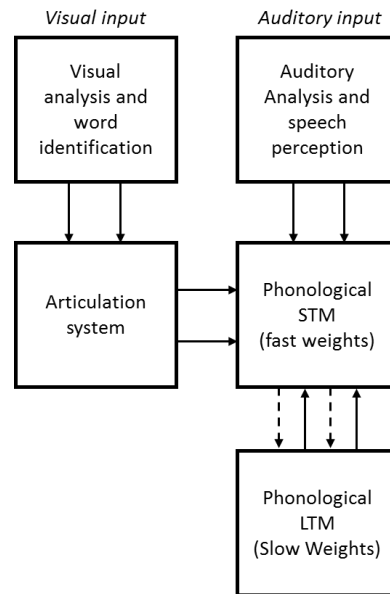


Figure 12. The phonological loop model proposed by Baddeley & Hitch (1974, adapted).

1.2.2.2.2 The Gupta model

Gupta (1996, 2003) proposed another model explaining the role of word repetition and immediate serial recall in word learning. Specifically, he proposed a 3-layer model with (i) the *semantic/context* layer where semantic information of the word is represented, (ii) the *phoneme* layer, where phonemes are represented (sublexical level), and (iii) the *phonological chunk* layer where word-forms are represented (**Figure 13**). These three layers are connected in such a way that a word-form in the phonological chunk is also represented in the semantic layer and, at the same time, can be produced by a specific sequence of phonemes represented in the phoneme layer. Specifically, sequences of sublexical elements are transduced into an internal word-form phonological representation. At the same time, this internal word-form representation can be transformed into a sequence of phonological output. These word-form phonological representations are subserved by the same shared semantic representations, a step that is accomplished via temporally learning in the STM. As opposed to Baddeley's model, where only STM is considered, long term memory is present in Gupta's model. He proposed that previous knowledge (long-term phonological knowledge) influences both nonword repetition and immediate serial recall facilitating the learning (Gathercole et al., 1991; Hulme et al., 1991).

Introduction

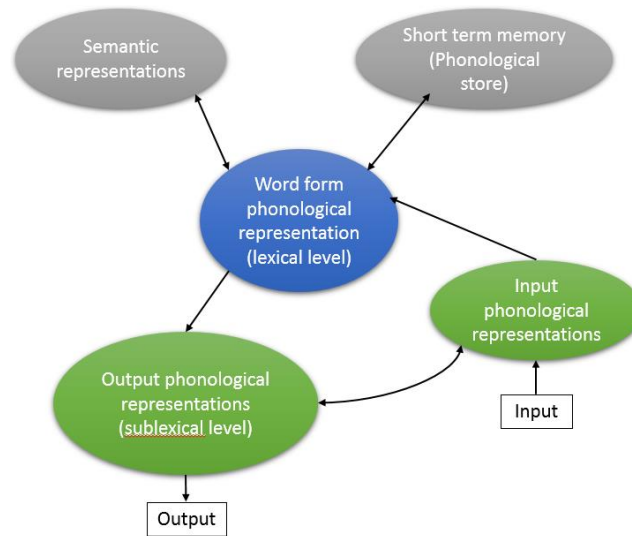


Figure 13. Representation of the Gupta model (2003, adapted).

1.2.2.2.3 The declarative/procedural model (DP)

A different model was introduced by Ullman (2001; 2004): the *DP* model. He proposed that two distinct types of long-term memory, declarative and procedural, support the mental lexicon and mental grammar. The declarative system may be responsible for acquiring initial knowledge due to its fast and flexible characteristics and its role in associative binding mechanisms. Furthermore, this type of knowledge or stored information can be consciously recollected. The DP model proposes that the declarative memory system regulates the mental lexicon, allowing it to store word-related knowledge such as meaning and word category, including abstract lexical representations. According to the model, declarative memory relies on MTL structures, including hippocampal regions, the entorhinal, perirhinal, and parahippocampal cortices (Squire & Knowlton, 2000).

In contrast, the procedural system is involved in storing new sensory-motor information, such as speech production. The procedural memory system may rely on basal-ganglia and frontal regions, including the supplementary motor area (SMA). These regions may underlie mental grammar, including syntax and all rule-based procedures that regulate the language and its hierarchical relations. Despite both systems, the declarative and procedural may have the same function: acquiring information, they can work in parallel, independently of one another, or even interact with each other, leading to an optimized functioning (Poldrack & Packard, 2003).

Overall, several models attempted to explain how memory networks interplay with language-related networks to facilitate learning. Despite differences in the models, it seems to be a

Introduction

consensus (i) on the involvement of STM in the initial stages of learning and (ii) on the interplay between existing memory traces and novel word-forms.

1.3 Electrophysiological correlates of word learning

The measurement of the brain's electrical activity allows the study of specific neural functions. We can measure *in vivo* brain activity at the neural population level with non-invasive surface EEG or with invasive *intracranial stereotaxic EEG recordings (sEEG)*. EEG and sEEG yield a continuous recording of brain activity from which we can obtain different information. In the following section, I describe some of the electrophysiological methods that I have been using during my Ph.D. Specifically, I focus on event-related potentials (ERPs) and brain oscillations.

1.3.1 Event-related potentials

The ERPs are obtained by averaging the activity recorded in response to a repeated external stimulus (Gazzaniga, 2014). This method is used in cognitive neuroscience to explore a large range of linguistic aspects such as lexical and syntactic processing in L1 and L2 (see Luck, 2005; Rodríguez-Fornells et al., 2009). Hereunder, I describe the most studied ERP component related to lexico-semantic processing, the N400 component.

1.3.1.1 The N400 effect

The first reported instance of the N400 component was published by Kutas and Hillyard (1980) in a study where participants listened to sentences with congruous or incongruous word endings. Semantically anomalous sentences elicited a negative deflection peaking at around 400 ms after word onset compared to standard endings sentences. Following this seminal work, the *N400 effect* has been observed for both written and auditory presented words (Kutas & Van Petten, 1988, for a review). For instance, it has been observed that low cloze probability words elicited a larger N400 than high close probability words (a measure of expectancy of words in a specific sentence position, Kutas & Hillyard, 1984). A reduction of the N400 amplitude is also observed using semantic priming paradigms (see Kutas & Federmier, 2011, for a review). Results converge in a decrease of N400 amplitude for semantically related primed words compared to unrelated primed words. Overall, there is a long tradition of studying the N400 effect as an index of lexico-semantic processing. The research on the *N400 effect* has been largely explored over the years with different experimental paradigms in both L1 and L2 (see a summary in **Table 1**).

Introduction

1.3.1.1.1 The N400 effect during contextual learning

After these initial experiments, several studies have used the N400 as an index of word learning in L1 during contextual learning. Perfetti and colleagues (2005) performed an experiment in which novel pseudo-words were presented together with matching definitions. After the training, they asked participants to make a judgment task about the relatedness or not of pairs of words while recorded EEG. The presented words could be pseudo-words, either previously trained or not, or real words. Furthermore, each pair of words could be related or not. They found that the N400 was more positive for trained pseudo-words than other word types over posterior locations during the first word presentation. For the second word, an N400 effect was found for trained and real words but not for untrained words over centro-parietal electrodes (**Figure 14A**). This result suggested that the N400 effect only occurred when participants successfully learned and integrated the meaning of new words.

In another study, Mestres-Missé and colleagues (2007) exposed participants to visually presented pseudo-words embedded in three meaningful high or low constrained context sentences. In the meaningful condition, the meaning of the word was consistent across sentences. Instead, in the low constrain condition, the word meaning was inconsistent across sentences. Results showed that participants could learn the meaning of the novel word in the meaningful condition after only three exposures. Importantly, by the third exposure, in this condition, the N400 to novel acquired words over centro-posterior regions was indistinguishable from the one to real words. On the contrary, novel words presented in low constrained context sentences showed a larger N400 than real words (**Figure 14B**). Interestingly, they found an N400 priming effect for newly learned words over frontal regions. This experiment showed that the N400 effect was sensitive to the fast acquisition of novel word meanings in L1 even with few words of exposure.

In the same line, Borovsky and colleagues (2010) recorded EEG while participants were visually exposed to known and unknown words in high or low constrain sentence contexts. After each word, they asked participants to judge the plausibility in the usage of the novel word. As expected, the results showed that the *N400 effect* during the exposure phase occurred over central electrodes within known words only. Interestingly, in the plausibility judgment task, the N400 was significantly higher for implausible compared to plausible judgments in all conditions (high-known, high-unknown, low-known) except for the low-unknown condition (**Figure 14C**). This effect was maximal over right central scalp locations. Overall, the findings confirmed the

Introduction

results of previous experiments and showed that the N400 effect occurred with only a single exposure to novel word-form in a meaningful context. Frishkoff and colleagues (2010) found a very similar result, who visually presented to participants novel words embedded in high or low constrained contexts and compared ERP responses to words with familiar (real) words before and after the training. They showed that both familiar and novel words trained in a high constrained context elicited N400 effects after the training. By contrast, low-constrain words did not show this effect.

Batterink & Neville (2011) performed an experiment in which the words were repeatedly presented ten times but embedded in short written stories instead of presenting the novel word-forms in isolated sentences. In the meaningful condition, the word meaning was consistent and deducible through the story, whereas no consistent meaning was available in the meaningless condition. They found that during the story presentation, meaningless words elicited a larger N400 than meaningful words. After the exposure, participants did a lexical decision task (to assess implicit learning) and a recognition task (to assess explicit learning). The results showed that while no N400 effect was observed in any of the conditions (meaningful nor meaningless) in the lexical decision task. In the recognition task, an N400 effect was found for meaningful words and also for real words over centro-parietal regions (**Figure 14D**). Based on these findings, the authors concluded that while explicit representations of the novel word-form emerged rapidly, implicit representations require further exposure or consolidation.

Finally, Bakker and colleagues (2015) went one step further and tested whether the N400 effect lasted 24 hours after the exposure to novel word-forms. In the training session, participants were exposed to novel word-forms and real words with their corresponding sentence definitions. Twenty-four hours later, participants were trained with a novel set of word-sentence associations and to real words not presented the day before. Consecutively ERP responses during a semantic priming task to old, new novel words and real words were obtained and analyzed. They found that novel words presented on the first day (novel remote) preceded by semantically related words showed similar responses to both recent and novel real words, that is, a decrease of the lexical N400 effect after consolidation (**Figure 14E**). Differentially, this effect was not found for recent novel words. Interestingly, while the N400 effect to recent novel words showed a fronto-central distribution, the N400 effect shifted to parietal scalp sites after consolidation.

Introduction

1.3.1.1.2 The N400 effect in associative word learning experiments

Other researchers used a different approach for testing the role of the N400 during meaning acquisition. Instead of using sentences, contexts, or definitions, they presented the novel word-associated with a specific picture. For instance, Angwin and colleagues (2014) asked participants to learn the name of aliens that were visually presented during a four-day training program. Half of the picture-words associations were accompanied by semantic information in which two adjectives related to the alien were provided (semantic condition). The other half of the picture-word associations were presented with two words with no semantic information (surnames of the aliens, non-semantic condition). They found better recall of the word-picture associations accompanied by semantic information than those without semantic information. On the fifth day, EEG was recorded while participants performed a picture-word judgment task. Results showed a higher N400 incongruency effect for the words in the semantic condition than for the non-semantic, suggesting a better integration of words with semantic information. In a similar direction, Bermúdez-Margaretto and colleagues (2018) exposed participants to novel auditory words alone or associated with real objects. They showed that when novel word-forms were associated with real objects (and thus when meaning was provided) showed a higher N400 effect over fronto-central electrodes in this condition compared with when words were not associated with a meaningful picture.

Overall, the studies reviewed above show that the N400 is a good marker of lexico-semantic integration of novel word-forms in L1 and that semantic information facilitates the integration of the novel word irrespectively of how the semantic information is provided to the learner.

Introduction

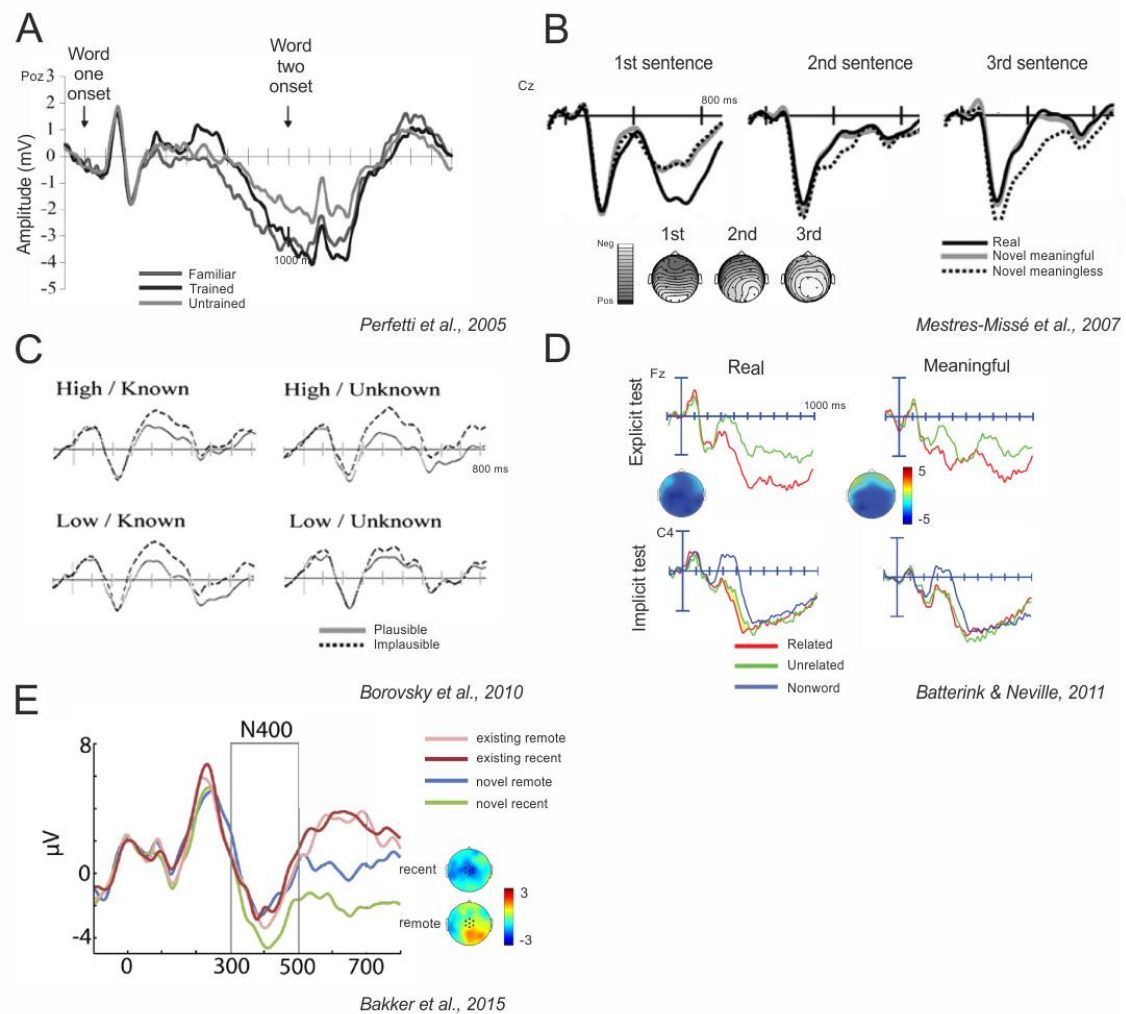


Figure 14. Selection of N400 experiments in L1. A. Grand average ERPs to familiar, trained words, and untrained words during the semantic decision task (from Perfetti et al., 2005). B. ERPs to real, novel meaningful and novel meaningless words during the contextual learning task (from Mestres-Missé et al., 2007). C. Results of Borovsky et al., 2010 to plausible and implausible word pairs for the high-known, high-unknown, low-known and low-unknown conditions during the judgment task (ERPs showed the average of right-medial electrodes). D. ERP responses from the experiment of Batterink & Neville (2011) during the explicit (recall task) and implicit tasks (lexical decision task) for related, unrelated words and non-words. E. Grand average ERPs to novel-remote, novel-recent, existing-remote and existing recent conditions over middle central electrodes (Bakker et al., 2015).

1.3.1.1.3 The N400 effect in second language learning

The N400 has also been proposed as an index word learning in L2. McLaughlin and colleagues (2001) were the first to demonstrate that the N400 effect was also sensitive to novel word learning in L2. Participants enrolled in a 1-year French course were assessed in their success in L2 three times along the course (after two weeks, in the middle, and at the end). The assessment consisted of recording the brain response during a priming task in which trained words (French) were presented with other French words that could be or not semantically related or with an

Introduction

unfamiliar pseudo-word. Furthermore, they compared the ERPs with a group of participants not enrolled in the course. In the first EEG session, learners showed larger negativity for unfamiliar pseudo-words compared to French words. However, no differences between semantically related or unrelated French words were found. Interestingly, by the third session, semantically unrelated word pairs elicited larger N400 than semantically related words (see **Figure 15A**). Conversely, no N400 effect was observed in non-learners in any of the sessions. The experiment evidenced that the integration of novel L2 words could also be tracked with electrophysiological recordings. In a similar vein, Soskey and colleagues (2016) taught English speakers' Spanish words over a semester of classes. During these periods, three EEG sessions were performed where participants were asked to perform a go/no-go lexical decision task. Results showed a larger N400 for native than L2 words in the first session. Importantly, this difference was progressively reduced along the semester of training over centro-parietal regions. By the third session, the N400 at frontal sites was very similar to an L1 response, indicating an integration of the novel words (see **Figure 15B**).

Other researchers studied the N400 effect in L2 in more controlled learning sessions (laboratory environment) and shortened language exposure. Stein and colleagues (2006) tested the effect of language exposure and familiarity of native English speakers to i) L1, ii) to new language trained during two learning sessions (L2, German), and iii) to an unknown language never trained (L3, Rumantsch). EEG was recorded at the end of both learning sessions in a task where participants had to indicate whether they knew or not the meaning of the presented word (that could be from L1, L2, or L3). They found that while the N400 decreased for learned L2 words in the second session compared with the first over parietal regions, no differences in the N400 were found between sessions for the non-trained L3 language (**Figure 15C**).

Yum and colleagues (2014) asked English speakers to learn Chinese words over 10-training-sessions in which four of them the EEG was recorded while participants performed a go/no-go semantic categorization task. Results showed that despite an overall improvement of all participants, individual differences in the rate of learning were observed. Furthermore, while fast learners exhibited an increase of N400 amplitude throughout the training sessions with an anterior distribution, this was not the case for slow learners (see **Figure 15D**).

More recently, Dittinger and colleagues (2016) asked participants to learn Thai words by auditorily presenting to them the novel words together with their corresponding picture while recording the EEG signal (during the learning task). Furthermore, they compared the

Introduction

performance of this task between musicians and non-musicians. The results showed an increase in the N400 amplitude during the exposure in musicians only (**Figure 15E**) over frontal regions. Interestingly, the initially frontal N400 rapidly shifted to parietal scalp sites in both groups. N400. Importantly, for the interpretation of the results, Thai words were composed of phonemes that do not exist in the native language of the participants (French). Hence, these results provided evidence of a fast integration of novel word forms and that musicians are better than non-musicians in building picture-word associations in a novel language containing unfamiliar phonemes. Dobel and colleagues (2009) used a similar word-picture paradigm but used pseudo-words with fricative phonemes instead. In line with previous experiments, participants showed a reduced N400 to novel word-forms compared to real words before the learning task. This difference was abolished after the learning indicating the lexico-semantic integration of the novel word forms.

Overall, the reviewed studies show that the N400 is also sensitive to word learning during L2 acquisition. Furthermore, the N400 effect occurs regardless of the learned language, and it is visible from a few minutes of exposure to the novel language to more extensive training programs.

Nonetheless, despite the vast literature suggesting that N400 modulations reflect the lexico-semantic integration of novel words, there is a lack of consensus regarding the direction of these modulations. While some studies have reported a decrease in N400 amplitude to words as a function learning (Stein et al., 2006; Dobel et al., 2009; Bakker et al., 2015; Stuellein et al., 2016), others have shown an increase (McLaughlin et al., 2004; Yum et al., 2014; Dittinger et al., 2016; Soskey et al., 2016). One crucial aspect that could explain this difference is **the amount of training** in the experiments. While some studies reported the effect after only a few exposures within the same training session (Perfetti et al., 2005; Mestres-Missé et al., 2007; Borovsky et al., 2010; Batterink & Neville, 2011; Dittinger et al., 2016; Stuellein et al., 2016) others reported differences within days (Stein et al., 2006; Frishkoff et al., 2010, Angwin et al., 2014; Bakker et al., 2015) or months (McLaughlin et al., 2004; Soskey et al., 2016). Another important issue that could explain those differences is **how the meaning is associated with the novel word**. Some studies used sentences with high or low meaning constrains to provide the meaning of the novel words (Mestres-Missé et al., 2007; Frishkoff et al., 2010; Batterink & Neville, 2011), while others used word definitions (Bakker et al., 2015) or word-picture associations (Dittinger et al., 2016; Dobel et al., 2009).

Introduction

| Reference | N° EEG sessions | Stimuli | L1/L2 | Amount of training | Task (learning assessment) | N400 modulation | Topographical distribution |
|----------------------------|-----------------|---|----------|-------------------------|-------------------------------------|--|----------------------------|
| Mc Laughlin et al., 2004 | 3 | Related or unrelated word pairs and pseudo-words | L1/L2 | 1 year | Lexical decision task | Increase of the N400 across sessions | Central |
| Perfetti et al., 2005 | 1 | Trained, untrained & familiar words with definitions | L1 | 45 min | Semantic decision task | Decrease of the N400 in trained words compared with untrained | Centro-parietal |
| Stein et al., 2006 | 2 | L1 familiar, L2 trained & L3 unknown words | L1/L2/L3 | 5 months | Recognition task | Decrease of the N400 for L2 trained words | Centro-parietal |
| Mestres-Missé et al., 2007 | 1 | Pseudo-words in meaning or meaningless sentence contexts | L1 | 3 exposures | Learning task | After training the N400 to meaningful words did not differ from real words | Fronto-central |
| Dobel et al., 2009 | 2 (MEG) | Identical or semantic related words associated with pictures | L2 | 5 training sessions | Categorization task | mN400 is reduced after training only for the learned words | Left temporal lobe |
| Midgley et al., 2009 | 1 | L2 words and their translation | L1/L2 | Beginners in L2 | Go/no-go lexical decision task | L1 words showed larger N400 than L2 words | Centro-parietal |
| Borovsky et al., 2010 | 1 | Pseudo-words & real words in high or low constrain sentence context | L1 | 1 exposure | Lexical decision task | The N400 was largest for low constraint pseudo-words | Fronto-central |
| Frishkoff et al., 2010 | 2 | Rare words trained in high or low constrains, rare words untrained & familiar untrained words | L1 | 2-day training sessions | 2-day training session | High reduction of N400 in rare words trained in high constrains but a low reduction in low constrain | Parietal distribution |
| Batterink & Neville, 2011 | 1 | Pseudo-words in meaningful & meaningless contexts (stories) | L1 | 10 exposures | Lexical decision task & recall task | Decrease of N400 across exposure | Left lateralized |
| Angwin et al., 2014 | 1 | Pseudo-word – picture association with 2 adjectives or 2 surnames | L1 | 4-day training sessions | Recall task & recognition task | Increased N400 in incongruent condition compared to congruent | Right lateralized |

Introduction

| Reference | N° EEG sessions | Stimuli | L1/L2 | Amount of training | Task (learning assessment) | N400 modulation | Topographical distribution |
|----------------------------------|-----------------|--|-------|----------------------|---|---|----------------------------|
| Yum et al., 2014 | 4 | L2 words | L1/L2 | 10 training sessions | Semantic categorization task | Increase of the N400 with training in fast learners only | Frontal distribution |
| Bakker et al., 2015 | 1 | Pseudo-words with definitions & real words trained or untrained | L1 | 2 training sessions | Recall & recognition task | Decreased N400 amplitude for trained words compared to untrained | Centro-parietal |
| Dittinger et al., 2016 | 1 | L2 words -picture pairs | L1/L2 | 20 exposures | Semantic & phonological categorization task | Increase of the N400 | Parietal distribution |
| Stuellein et al., 2016 | 1 | Trained or untrained L1 words associated with few or many semantically related words | L1 | 4 training sessions | Recognition task | Decreased of the N400 for trained words and larger N400 of words with many semantics | Fronto-central |
| Soskey et al., 2016 | 3 | Novel L2 words and their translation | L1/L2 | 6 months | Go/no-go lexical decision task | Increase of the N400 to novel words across sessions | Centro-parietal |
| Qi et al., 2017 | 1 | Acceptable or unacceptable L1 sentences | L1/L2 | 3 training sessions | Acceptability judgment task | Individuals with a larger N400 effect in L1 are better in L2 language learning | Centro-parietal |
| Bermúdez-Margaretto et al., 2018 | 1 | Pseudo-words associated with pictures | L1 | 6 exposures (blocks) | Lexical decision task | Decrease of N400 amplitude after training when meaning but not in a meaningless context | Fronto-central |

Table 1. Summary of ERP studies on word learning. L1 = first language, L2, second language.

Introduction

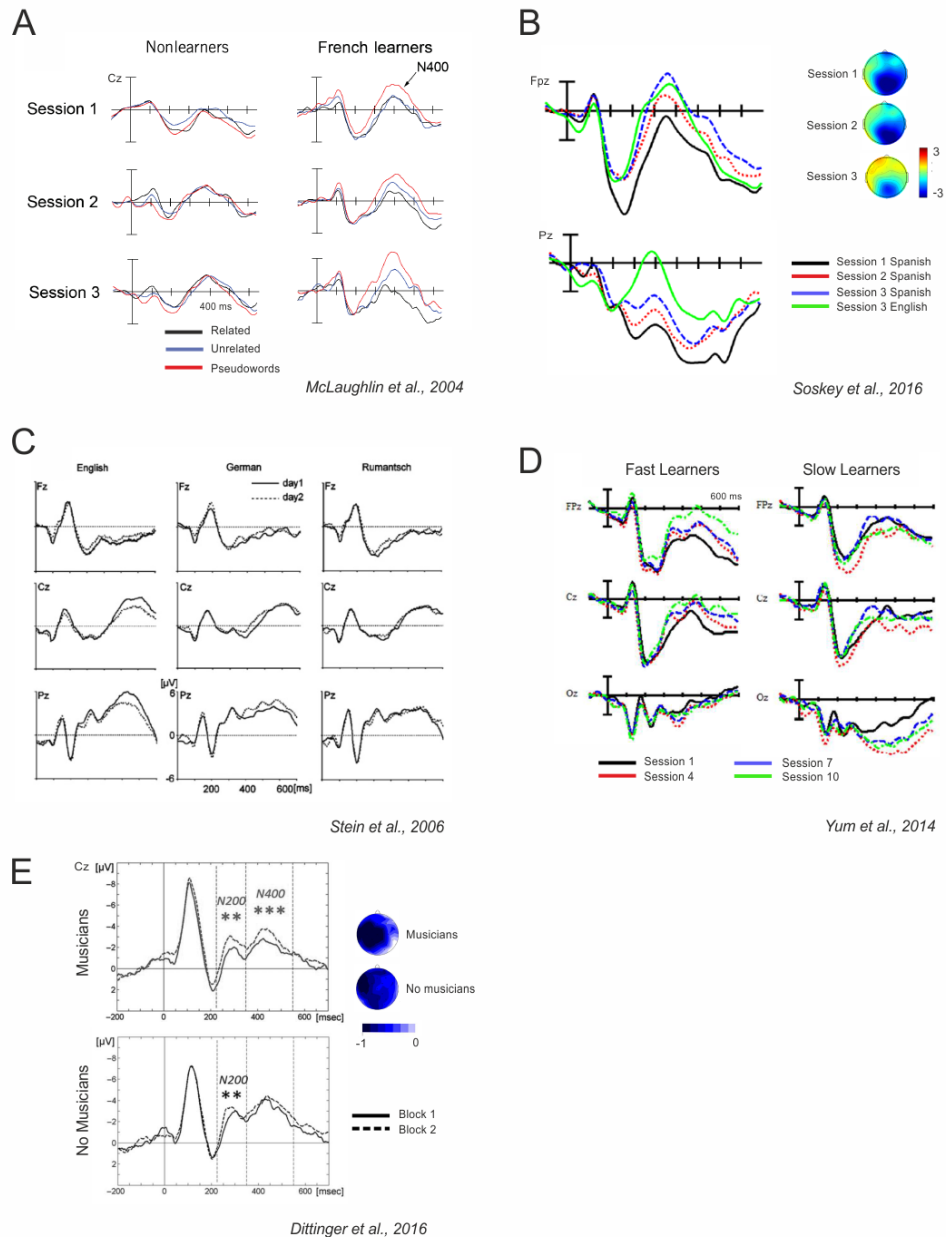


Figure 15. Selection of N400 experiments in L1. **A.** Grand average ERPs response during the three testing sessions for the non-learners and French learners (McLaughlin et al., 2004). **B.** Grand average and voltage map to L2 during the three testing sessions and to L1 during the last session. **C.** ERP response found in Stein et al., 2006 to English, German and Rumantsch. **D.** ERPs response through the training sessions for slow and fast learners. **E.** Grand average of ERPs response from the study of Ditinger et al., 2016 during the learning phase for musicians and non-musicians.

Notably, the **topographical distribution of the N400** reported across experiments also differs. While some studies showed a modulation of the N400 over prefrontal regions (Stein et al., 2006; Soskey et al., 2016), others reported modulations over centro-parietal regions. These

Introduction

differences could be explained in the context of the time-dependent hypothesis that proposes an increase of activation during the very initial learning period (Poldrack, 2000; Raichle et al., 1994) followed by a decrease in activation that may reflect an optimization process (Chein & Schneider, 2005; 2012). Furthermore, the different designs and materials used in the literature may target different learning stages. They may also involve different cognitive processes such as cognitive control and monitoring processes, typically related to prefrontal activation, but also lexical access, semantic integration, and semantic priming (Lau, Phillips & Poeppel, 2008), which require the involvement of language-related regions.

1.3.1.1.4 The N400 in infants and children

The N400 component has not only been investigated in adults but also in infants, where an N400-like potential has been observed. Some researchers have used the classical object-word learning paradigm to explore the modulation of the N400 in infants. Friedrich and Friederici (2008) exposed 14-month old infants to novel object-words pairs. They demonstrated that at this age, with only four exposures to word-picture pairs, infants were already able to associate a meaning with a novel object as reflected by a modulation in the N400-like component. Torkildsen and colleagues (2008) used a similar paradigm and found that infants with a wide range of vocabulary showed an N400 incongruity effect on incorrectly paired word-objects, whereas infants with lower vocabulary did not. Borgström and colleagues (2015) tested infants between 20 and 24 months old in a similar paradigm and demonstrated a correlation between their productive vocabulary and the N400 effect for the pseudowords but not the real words. Junge and colleagues (2012) asked 9-month-olds' ability to learn basic level picture-word associations. Importantly, during the training, half of the children saw constant images for each word, whereas the other half saw different images of the same object (e.g., six types of different words). In the consecutive test, participants were exposed to word-picture associations that could be congruent or incongruent. They found that both groups of participants showed modulation of the N400 during the test phase when presented with incongruent object-word associations suggesting that infants have identified the concept of the picture in both cases. A different paradigm was used by Friedrich and Friederici (2011). They used a cross-modal learning paradigm to demonstrate that 6-month-olds can associate novel words with novel objects using this paradigm and found an N400 priming effect during the training phase.

Importantly, although many experiments studying modulations of N400 have been conducted with adults and infants, very few have been done with children. In a recent study, Dittinger and

Introduction

colleagues (2017) exposed 8-12-year-old children to Thai words associated with pictures and found an increase of N400 amplitude through the learning task. Overall, the literature with infants and children points towards the N400 as a neural marker of building meaningful object representations for novel objects (Friedrich & Friederici, 2017), familiar objects (Friedrich & Friederici, 2004), or both (Torkildsen et al., 2008). *These data converge with the adult literature and also provide evidence that word learning can be tracked with electrophysiological measures at a very young age.*

1.3.1.1.5 The N400 effect in the context of memory

Finally, the N400 has also been studied in the context of memory. Several studies have proposed that the N400 can be modulated by familiarity (Rugg & Curran, 2007); that is, the fact that a stimulus is previously presented. These studies showed a decrease in N400 amplitude once words become familiar to the learners. These studies support the *dual-process theory of memory*. This model proposed by Curran (2000) proposed that recognition of word-forms or images (memory recognition) can activate conceptual representations gathered in memory via two different processes: *familiarity* and *recollection*. The concept of *Familiarity* refers to the situation of recognizing but not remembering something (James, 2007). *Recollection* refers to the retrieval of details regarding how the knowledge was acquired. While familiarity is considered as a relatively fast and automatic process, recollection, by contrast, is conceived as a slower process that requires conscious access to information about the context. Evidence for the dual-process model mainly comes from studies using ERPs that have identified distinct components for the two processes. During recognition memory tests, a negative component peaking between 400 and 800 ms after stimulus onset was reported with novel words showing larger negativity than old words. This effect was maximal over parietal regions and therefore called the “parietal old/new effect” related to recollection. Besides, a frontal N400-like component (FN400) has been related to familiarity (Curran, 2000). Despite the literature talking about the model (Paller, Voss & Boehm, 2007; Rugg & Curran, 2007), some researchers claimed that both FN400 and N400 reflect the same semantic processing during recognition test (Voss & Federmeier, 2011). In a similar direction, a recent study showed that the N400 is not modulated by familiarity or repeated exposure, but only when words were presented in a lexico-semantic context (Bermúdez-Margaretto et al., 2018). This arises because several encounters with the word-object referent are usually required to consolidate the word form in the lexico-semantic network. *Overall, the debate is still open, and further research needs to be done to resolve this issue.*

Introduction

1.3.1.2 Electrophysiological studies of speech segmentation based on SL

Some researchers have used electrophysiological measures to investigate the neural bases of SL. Sanders and colleagues (2002) were the first to use the SL paradigm while recording EEG. In their experiment, participants were exposed to a 14-minute artificial language stream composed of six tri-syllabic words with the TP between words as the only cue to word boundaries. ERP responses recorded during stream exposure showed larger N1 responses to word onset post-exposure than pre-exposure. The N1 component was previously observed in response to the initial syllable of a word, and thus, authors conclude the emergence of N1 after training indexes word segmentation. The same study also revealed N400 effects in response to learned words following; thus, the same pattern as classical N400 word learning experiments. In a subsequent study with real sentences, the same authors (2003) found similar effects on the N1 component at the initial syllable of the word even when initial and medial syllables had the same intensity and duration. Neither semantic nor syntactic information modulated the N1 effect. Thus, they proposed that this effect reflected selective attention to word onset rather than reflexive attention. A similar result was found in a study by Astheimer and Sanders (2011). They showed that selective attention to word onset occurred when onsets cannot be predicted from the context. It is worth mentioning that in these studies, participants were explicitly instructed to learn the novel words and that the attentional mechanisms involved could then be different when learning is implicit (Toro-Soto & Trobalón, 2005). It is possible that considering this paradigm, the segmentation process may be driven by explicit instructions to search for newly learned words. Indeed, a recent study has demonstrated that explicit instructions can attenuate implicit learning (Fletcher et al., 2005).

In another experiment, Cunillera and colleagues (2006) recorded ERPs while participants listened to an artificial speech stream composed of four tri-syllabic words. Word boundaries were marked by statistical cues alone or by statistical and stress cues. The control condition consisted of randomly generated combinations of the same syllables used to create the words presented in the experimental condition. The authors found differences in the N1 component between the learning and random conditions for the stressed, but not for the unstressed streams (**Figure 16**). They also found effects on the N400 component, as in Sanders et al.'s experiment (2002). However, in Cunillera's experiment, the topography had a frontocentral distribution, while Sanders and colleagues (2002) showed a more parietal distribution. These differences were argued to be due to different experimental procedures. While the ERP analysis in Sanders's study focused on comparing brain responses collected before and after the training,

Introduction

Cunillera and colleagues focused on the on-line process. Despite the topographical differences, the N400 seems to be an index of initial traces of lexical acquisition in both cases.

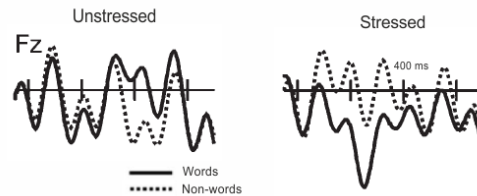


Figure 16. Grand average potentials from Cunillera et al., 2006 for words and non-words in the stressed and unstressed conditions during the on-line speech segmentation task.

De Diego-Balaguer and colleagues (2007) tested whether participants were sensitive to structural rules embedded in the words. They created an SL paradigm with non-adjacent statistical dependencies between syllables (with pseudowords with a fixed initial and ending syllable regardless of the middle syllable, e.g., **paliku** or **paseku**). They analyzed the ERP responses to the tri-syllabic pseudowords along the four minutes of exposure. Electrophysiological results showed a decrease of the N400 over central electrodes and a development of a frontal positivity (P200) through the task (as rule-learning was acquired, **Figure 17**). They also found that once the rule was learned, words violating the rule generated a frontal N400, similar to the lexico-semantic effects of the N400 seen in word learning experiments, indicating a lexicalization of the novel words.

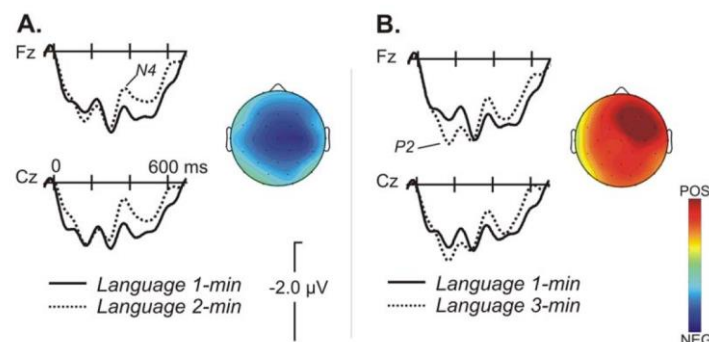


Figure 17. Grand average ERPs and topographical distribution of results from de Diego-Balaguer et al., 2007 for language novel words contained in the streams comparing the first and second minutes of exposure (A) and comparing the third and first minute (B).

Abla, Katahira & Okanoya (2008) explored whether modulations of the N400 during a segmentation task occurred with both speech stimuli and non-linguistic tones instead of words as in Saffran and colleagues (1999). They divided participants into three different groups

Introduction

according to their behavioural performance at the test and found that both the N1 and the N400 decreased through learning sessions in high and middle learners but not in low learners. They also replicated the experiment with visual stimuli (Abla & Okanoya, 2009) and obtained similar results. These findings suggest that SL is not specific to language but instead occurred regardless of the stimuli presented.

Although speech segmentation and meaning acquisition have been largely investigated separately, a recent computational model from Räsänen and Rasilo (2015) suggests that word segmentation may be facilitated when segmentation and meaning acquisition occur simultaneously (see section 1.2.2). François and colleagues (2017b) tested this hypothesis by exposing participants to a continuous stream of non-words associated or not with a visual referent (object). The authors found that words were better extracted from the stream when objects were consistently associated with the pseudo-words. Interestingly, they found a larger N400 amplitude for the consistent condition than for the inconsistent condition, indicating a better mapping of word-forms onto the lexico-semantic network.

The body of literature described so far suggests that humans are highly sensitive to the statistical properties of the environment and that this sensitivity plays an important role in language learning (see Siegelman, 2020, for a review). Furthermore, they also indicate that individual differences in sensitivity to SL can give rise to individual differences in language acquisition and comprehension (Siegelman et al., 2017; François et al., 2017a).

1.3.2 Brain oscillations

As mentioned in the introduction of this section, we can explore the properties of brain oscillations with EEG recordings (see the box “Brain waves properties”). In the brain, groups of neurons act together, firing in a synchronous way to provide a specific function. These repetitive patterns of neural activity are called **neural oscillators**. Oscillations play a role in various brain operations such as stimulus processing (Lakatos et al., 2005), cognitive control (Canolty et al., 2006) and are involved in regulating the flow of sensory information. However, cognitive tasks are not restricted to a single brain region but usually require the involvement of different brain regions instead. The flow of information between those regions generates a synchronized activity among brain regions (**brain oscillations**), giving rise to a large electric field that can be recorded using EEG and MEG (Hämäläinen et al., 1993; Buzsáki & Wang, 2012). By characterizing

Introduction

the amplitude and phase of the oscillations through brain connectivity measures, we can study how brain regions communicate with each other.

Significantly, rhythmic external stimuli can modulate the frequency and timing of brain oscillations. The brain's temporal dynamics become synchronized to the external pattern, making the oscillations increase or decrease in amplitude when a stimulus is expected or not. This mechanism facilitates the prediction of certain stimuli. When a stimulus appears periodically, the occurrence of that stimulus can be predicted. Hence, it has been proposed that the temporal alignment of the neural oscillations with a rhythmic stimulus facilitates the processing of such stimulus by allowing a faster response than an unexpected one (Mathewson et al., 2010).

1.3.2.1 Brain oscillations and speech perception

Recently, research has focused on how the brain transforms speech sound into meaningful units. Luo and Poeppel (2007) exposed participants to spoken sentences while recorded MEG to show that sentences could be distinguished from each other only by analyzing the patterns of auditory brain activity. This result occurred due to phase entrainment across trials between brain activity and the external stimulus (see **Figure 18**). Further, some studies have shown that a better synchronization with external stimuli is linked to better speech comprehension (Pelle et al., 2013) and better stimulus-response (Assaneo & Poeppel, 2018). Importantly, when neural tracking is not possible due to degraded stimuli, speech perception and comprehension are impoverished (Ahissar et al., 2001; Luo & Poeppel, 2007). *Despite these findings, it is not yet clear how the synchronization between brain oscillations and speech occurs.*

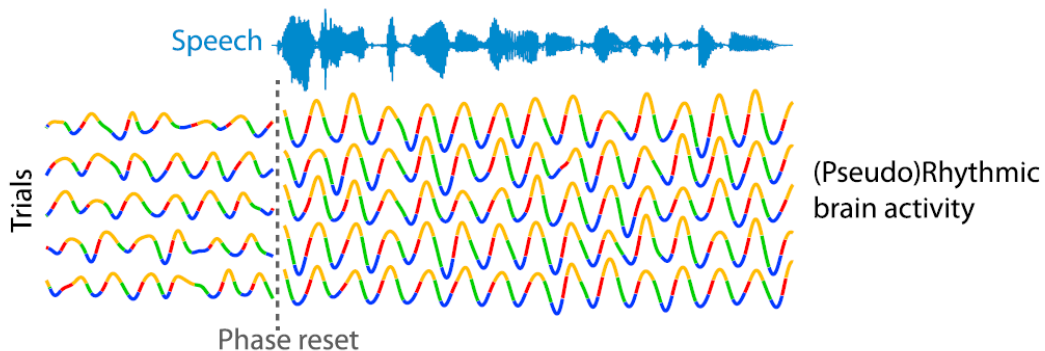


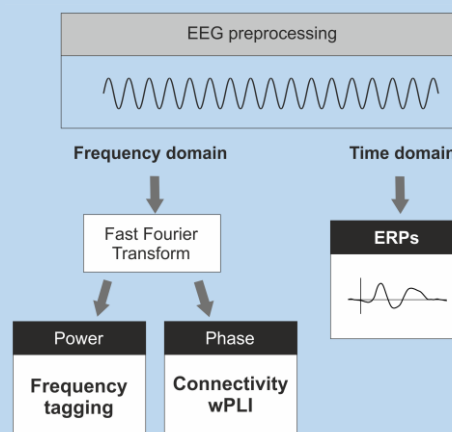
Figure 18. Representation of the neural synchronization of brain signals to a speech stream (Obleser & Kayser, 2019). The phase entrainment between brain activity and the external stimulus is represented by a consistent alignment of the phase (color-coded) across the trials of the stimulus. A possible mechanism could be a phase reset of the brain activity driven by the stimulus onset.

Introduction

Human speech is a complex auditory signal that contains rich dynamics in the amplitude and frequency domain (Zeng et al., 2005; Smith et al., 2002). Natural speech occurs at multiple time scales with close resemblances among languages. Phonemes, on average, appear about every 25 ms (~40Hz, gamma range), syllables have a duration of around 200 ms (~4-5 Hz, theta range), and words appear about every 700 ms (~1-2 Hz, delta range, Greenberg & Arai, 2004; Ding et al., 2017). Therefore, most of the relevant speech signal oscillates between theta (4-8 Hz) and delta bands (0.5-3 Hz; Hickok & Poeppel, 2007; Ding et al., 2016). The fact that speech signal oscillates within a specific frequency range may help to predict the onset of the syllables and, by extension, facilitates the speech segmentation process (Giraud & Poeppel, 2012). This hypothesis is supported by studies showing that the changes in oscillations are closely related to the onsets of syllables (Stevens, 2002).

Brain waves properties

Brain signals are electrical waves, and therefore, the same properties that describe a wave can be used to describe them. There are three properties of waves that allow to define them: frequency, power, and phase. *Frequency* answers the question “how fast is the wave?” and refers to the number of cycles that occur over 1 second (measured in Hz). *Power* answers the question “how strong is the wave?” and is defined by the amplitude of the wave (power = amplitude²). Finally, *phase* answers the question “At which point in the cycle is the wave?”. By analyzing their specific properties, we can obtain different information about the behavior of the wave. We can study how amplitude of the signal change along the time domain (ERPs), along the frequency domain (frequency tagging) or both (time frequency). We can also use the phase properties of the wave to evaluate how well two signals are synchronized (connectivity analysis).



When the sound enters through the ear, it reaches the cochlea, a structure that mechanically decomposes the auditory signal into different frequencies. The hierarchical decomposition of

Introduction

the acoustic signal is sent to the auditory cortex, where information is hierarchically organized (tonotopic organization; Lakatos et al., 2005). This hierarchical structure has been shown at oscillation level where delta amplitude (0.5-3 Hz) modulates theta amplitude (4-8 Hz) (that would correspond to phonemic-syllabic interaction), and theta phase modulates gamma (30-50 Hz; that would correspond to auditory cortex integration; Giraud & Poeppel, 2012; Ghitza et al., 2013, see **Figure 19**). This mechanism has been proposed to ease the tracking of speech envelope, hence facilitating speech comprehension (Doelling et al., 2014).

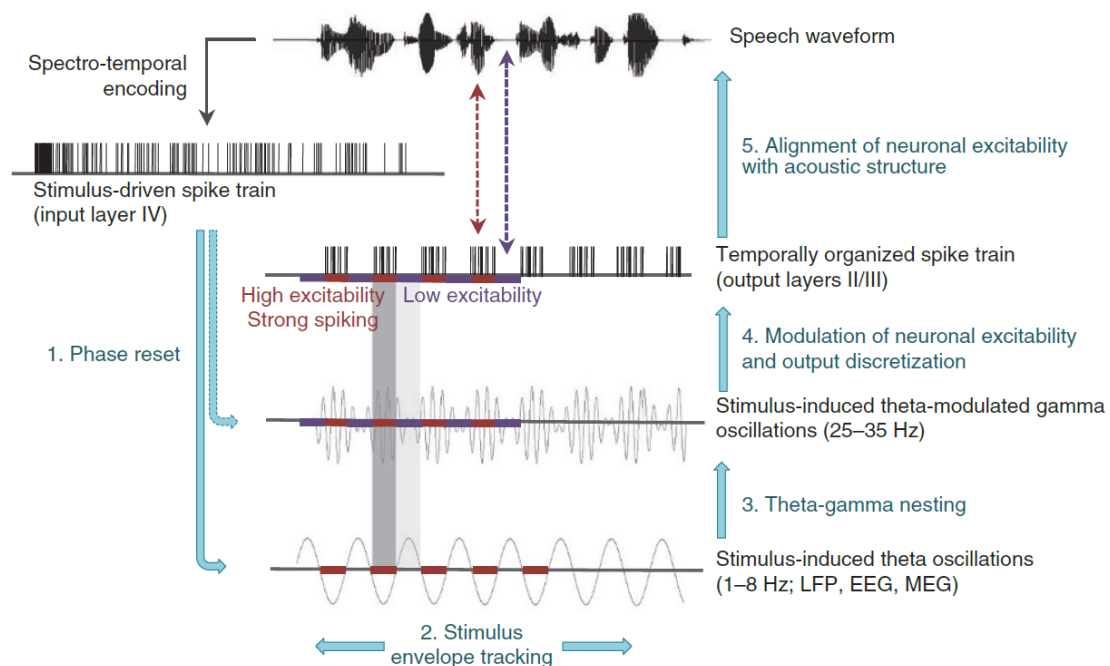


Figure 19. Oscillatory activity in speech processing from Giraud and Poeppel (2012). Speech input is first encoded in the primary auditory cortex, in which specific cortical layers respond to specific stimulus frequency. Response onset elicits a reset in theta oscillations in superficial layers (step 1). After the reset, theta oscillations track the speech envelope (step 2). Theta reset induces a reset in gamma activity and a subsequent theta-gamma coupling (step 3). Then, gamma activity controls the excitability of neurons generating feedforward signals to higher cortical areas (step 4). Finally, the neuronal excitability phase aligns to speech structure (step 5).

1.3.2.2 Oscillations and attention: can we control the neural synchronization?

It has been proposed that neuronal synchronization to an external stimulus can be modulated by two possible mechanisms. In the *bottom-up* or *stimulus-driven* approach, the external stimulus's salience drives the automatic synchronization of the neuronal oscillations. While in the *top-down* or *goal-directed* approach, attention is voluntarily directed to a specific stimulus.

Imagine a busy thoroughfare with car horns, idle chatter, emergency sirens where you are trying to have an important phone call. You may actively try to focus on the conversation while trying

Introduction

to ignore any other sounds. This phenomenon is called *the cocktail party effect* (Cherry, 1953), and since it was described, neuroscientists have tried to understand the neural mechanisms beyond this phenomenon. The most accepted hypothesis is that attention can be voluntarily directed toward the stimulus (*bottom-up attention*; Corbetta & Shulman, 2002). Ding and Simon (2012) asked participants to selectively listen to one of two competing speakers while they recorded MEG. They showed brain alignment to the attended speaker but not the unattended one. In a similar vein, studies with animals also showed that non-attended stimulus does not induce alignment in the respective tonotopic regions (Lakatos et al., 2013). This finding has been consistently replicated in the literature (Chait et al., 2010; Southwell et al., 2017;). Furthermore, sensory processing is also modulated by prior knowledge or experiences (Sohoglu et al., 2012). Those findings clearly favor the hypothesis of top-down attentional control during speech processing.

1.3.2.3 Origin of the oscillations: ERPs vs. oscillations approach

One of the main critiques of the oscillatory approach (e.g., frequency tagging measure) is that the synchronization observed in electrophysiological data could be explained by concatenated evoked responses (Capilla et al., 2011). We just described above that when a participant listened to a rhythmic stimulus, this rhythm was observed in the neural oscillations. However, intrinsic oscillations are ubiquitous in the brain and occur in the absence of exogenous stimulation (Obleser & Kayser, 2019). The question then is: “is the synchronized activity to a stimulus arising from the direct induction of the external stimuli?” that is, does the peak in EEG data represents neural entrainment or rather the synchronization of the ERPs produced after the presentation of the repetitive stimulus (Zoefel et al., 2018)? This dilemma arises because evoked responses are highly phase-locked to stimulus presentation (Van Diepen, Foxe & Mazaheri, 2019), thus over-estimating the oscillations at the stimulus frequency. Some researchers have tried to discern these issues by using indirect measures such as sub-threshold stimulation (Oever et al., 2017) and found that oscillation activity still arises in this condition. Others studied the oscillation activity while after the stimulus presentation based on the idea that an oscillator would continue for a while after the stimulus presentation (Lakatos et al., 2013). In a recent paper, Doelling and colleagues (2019) theoretically tested this hypothesis by creating two models that generated a neural response with a fixed time lag (evoked model) or at the frequency of the stimulus (oscillator model) and observed how phase varies as a function of frequencies. They found that while in the oscillation model, the phase remained stable across frequencies, in the evoked model, the phase did not, thus arguing in favor of the oscillatory

Introduction

model. Overall, despite some controversy in the field, the evidence seems to favor the oscillation hypothesis.

1.3.2.4 Methods in oscillations analyses

1.3.2.4.1 Frequency tagging analysis

One way of analyzing brain oscillations is to use steady-state potentials, that is, to use the property of the brain to adapt its oscillatory activity to the rhythm of a visual or auditory stimulus presented periodically (Regan, 1982; Picton et al., 2003). In this type of analysis, we can use the Fast Fourier Transform (FFT) to decompose the raw EEG data into “basic” waves of different frequencies. Indeed, the Fourier theorem proposes that complex waves can be explained by basic sinewaves of different frequencies (Cohen, 2014). After applying the FFT, the power values obtained provide information about the strength of the neural response. It is also possible to compute inter-trial coherence (ITC) to obtain the phase values. The ITC gives information about the precision of the neural response by calculating the phase values in every trial and comparing it across conditions (Wöstmann et al., 2017). Both measures, raw power, and ITC have been used in frequency tagging analysis applied to SL.

Buiatti and colleagues (2009) used FFT to quantify the neural synchronization to syllable and word frequencies during a speech segmentation task based on SL. Participants were exposed to statistically structured speech streams composed of tri-syllabic pseudo-words of equal duration i) with and ii) without boundary cues or to random streams. They found a significant increase in power at word frequency when participants were exposed to the structured stream (when the words could be extracted) but not during the exposure to random stream (see **Figure 20**). In the random stream, they found a peak of power at the frequency of the syllables only. Interestingly, they found that neural synchronization to word frequency at the structured stream positively correlated with the percentage of correctly recognized words. These results suggested that during word learning, single syllables are bound together to recognize word units.

Introduction

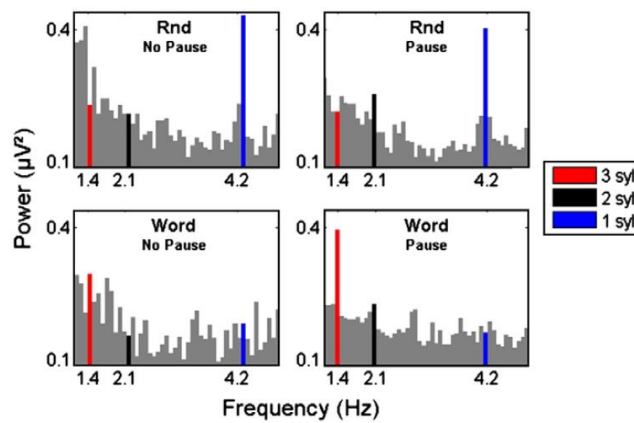


Figure 20. Results from Buiatti and colleagues (2009). The figure shows the EEG averaged power spectrum during exposure to the streams (random and learning streams with or without pauses). The figure shows a clear peak at syllable frequency in the random condition, whereas a peak to three syllables (word frequency) appears in the word condition.

In another study, Batterink and Paller (2017) used ITC to test whether the synchronization to the stimuli occurred gradually through participants' exposure to a speech stream composed of tri-syllabic pseudowords. They found that the ratio of word frequency to syllable frequency brain response (word learning index, $WLI = \text{ITC of word frequency} / \text{ITC of syllable frequency}$) increased across blocks for the structured stream (**Figure 21**). This effect was driven by a progressive increase of ITC at word frequency and a progressive decrease at syllable frequency. They also observed a positive correlation between the WLI and the performance in the behavioral recognition task.

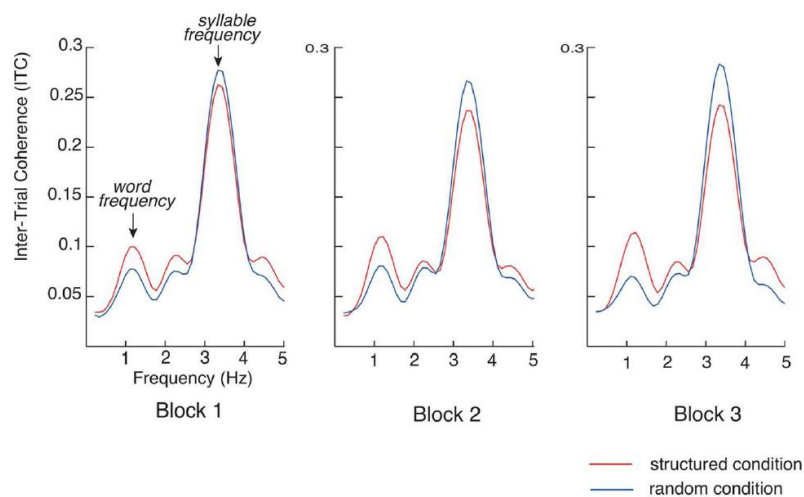


Figure 21. EEG results from Batterink and Paller 2017, along with the three learning blocks for structured and random streams.

Introduction

Nozaradan and colleagues (2011) used a frequency tagging approach to test whether neural synchronization also occurred to beat and meter in musical stimuli. They exposed participants to rhythmic stimuli and asked them to imagine the meter as being binary or ternary. They found that the brain signal oscillated at the specific imaginary frequencies (Nozaradan et al., 2011, see **Figure 22**). In another study, Nozaradan and collaborators (2017) showed that when introducing a contrast tone among a repetitive pure tone stimulus (e.g., AAAAB, where A & B are pure tones), a frequency band at the exact frequency of the contrast appeared on the brain signal.

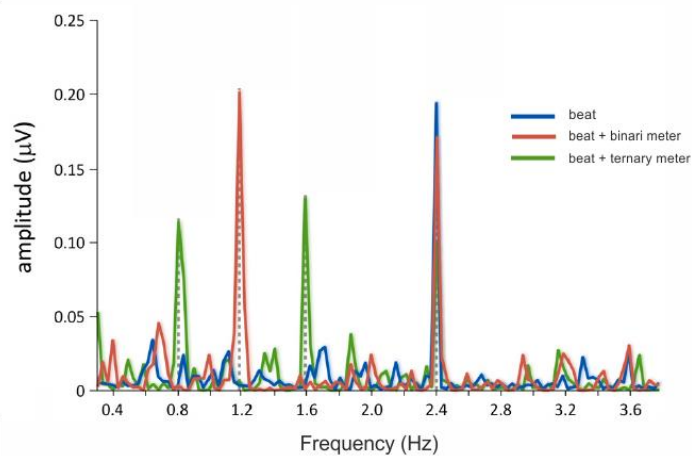


Figure 22. EEG responses from Nozaradan et al., 2011. Showing the brain response to the beat alone (blue), to the beat plus an imaginary **binary** meter (red), and to the beat plus an imaginary **ternary** meter.

Overall, this body of evidence shows that cortical regions are sensitive to regularities in the environment for different sensory modalities (auditory and visual; see Nozaradan et al., 2014, for a review). Further, some studies have also shown that cortical oscillatory activity is sensitive to different hierarchical levels of information such as sentences (Ding et al., 2016), prosodic cues (Elmer et al., in prep), or the presence of higher and lower TP (e.g., the co-occurrence of syllables in tri-syllabic words and that the same syllable could appear more than one word; Ordin et al., 2020). Finally, the literature reviewed also shows that frequency tagging analysis is a useful tool to study the synchronization of brain activity to rhythmic stimuli.

1.3.2.4.2 Connectivity analyses

The oscillatory activity also provides useful insights about the communication between brain regions that are adjacent or not (Bastos & Schoffelen, 2016). This approach is based on the premise that when two regions communicate, they oscillate at the same frequency (i.e., the

Introduction

fluctuations of local neuronal ensembles rhythmically synchronize in both regions; Buzsáki & Wang, 2012).

However, some limitations arise from recording methods such as the typical scalp EEG. For instance, one of the main problems of connectivity measures is the volume conduction problem. It is due to the large distances between the sensors and the neural sources. The current flow in the surroundings of a neuronal source causes one recording channel to pick-up activity from multiple neuronal generators (Bastos & Schoffelen, 2016). This confound makes it challenging to say whether the activity collected by an electrode reflects the activity only of the below cortical region or instead reflects also surrounding brain regions' activity. Another important issue in connectivity is the difficulty of distinguishing direct from indirect interactions, which means that we do not have the certainty that the interaction between two regions is direct or may come from a third source.

Despite the limitations, several metrics can be used for functional connectivity. The simplest way of measuring the correlation between two variables is the Pearson correlation coefficient. However, this measure does not take into account the temporal structure of the data. Thus, the equivalent measure for time series data is the cross-correlation. Though, time-series data is often represented in the frequency domain instead of the time domain. Hence, the **coherence coefficient** computes a cross-correlation in the frequency domain. This measure computes the amount of variance in one signal that can be explained by the other signal. Thus, coherence is used to determine if two or more sensors or brain regions have similar neuronal oscillatory activity. It quantifies the frequency and the amplitude of the brain oscillatory activity by estimating phase and amplitude consistency between signals from different electrodes/regions and various frequency bands. It is based on the idea that if two signals are "in phase", their amplitude will add, whereas if the signals are "out phase", the amplitude will subtract. The coherence value of 1 means that data is identical in both electrodes/regions, while 0 indicates no similarity between signals. Importantly, coherence does not solve the volume conduction problem mentioned above. An alternative, then, is the use of the imaginary part of the coherence. This measure projects the complex-valued coherence onto the imaginary axes, removing instantaneous interaction derived from the field spread.

Another approach for computing connectivity is to use phase synchrony. Phase synchrony measures how stable the phase difference varies over a short period. Phase relationship can be measured as to how the phase changes across trials within one electrode (inter-trial phase

Introduction

coherence, ITPC) or comparing how the phase relationship varies across trials between pairs of electrodes or regions (phase-locking value). The phase lag index (PLI) is another measure of phase synchronization that estimates whether consistent non-zero phase lag exists between two signals (Stam et al., 2007; Vinck et al., 2011; Hardmeier et al., 2014). It is computed by averaging the sign of the phase difference of each per observation (Stam et al., 2007). This measure is motivated by the fact that non-zero phase differences cannot be caused by the field spread. It has the advantage over other measures (i.e., imaginary coherence) of being less influenced by phase delays. However, the PLI overestimates the population statistics for a finite sample size. With the **weighted PLI (wPLI)**, the contribution of the observed phase leads and lags is weighted by the magnitude of the imaginary component of the cross-spectrum (Vinck, 2011). This measure also accounts for sample-size bias by squaring the PLI (Vinck et al., 2011). For the advantages described above, the wPLI has been reliably used as a connectivity measure in several experiments (Reiterer et al., 2011; Ortiz et al., 2012; Chaturvedi et al., 2019) and thus, has also been used in the present work.

Chapter 2

Research aims

Chapter 2 Research aims

As reviewed in the introduction, language acquisition is a complex process that involves several learning steps: isolating the words from the fluent speech, associating meaning to the novel word, and integrating the word-meaning into the current lexicon. Those steps require the involvement of cognitive processes supported by cortical and subcortical structures. Although much research has addressed how the language acquisition process occurs, there are still many unresolved questions about how the different learning steps temporally co-occur and which brain regions contribute to these steps. ***Therefore, this Ph. D. thesis aims to extend the current knowledge on the neural bases of language learning mechanisms in adults and children with behavioural and electrophysiological measures.***

Specifically, we aimed to:

- i) understand better how novel word-forms are learned and which is the role of domain-general mechanisms (e.g., cognitive control or attention) in this process;
- ii) understand better how multimodal information is used to concurrently isolate and associate meaning to novel words;
- iii) directly evaluate the role of the hippocampus in the speech segmentation process.

To address these questions, we designed three studies that provide a comprehensive picture of behavioral and electrophysiological brain mechanisms involved in language learning in adults and children. Thus, although different in terms of paradigms and populations, the three studies share a common focus on human language learning.

2.1 Study I: The interplay between domain-general and domain-specific mechanisms during novel word learning

In the first study, we studied the neural dynamics associated with the encoding and later consolidation of novel object-words pairs. Specifically, we aimed to understand better how novel words are integrated into the mental lexicon in the specific situation where existing semantic knowledge cannot be used. With these aims, we designed two ERP experiments. In *Experiment 1*, twenty-five participants learned novel object-word pairs during a 5-day training

Research aims

program. We assessed immediate recognition and recall of the new associations during each training day and a 4-month follow-up. We recorded EEG during the first and last training days to track the brain responses to novel words from the early stages of learning up to the later stages after consolidation had taken place. An innovative aspect of this study was that participants had to associate new words with unknown objects. Thus, the participants could not use pre-existing lexical memory traces to facilitate the learning (Angwin et al., 2014).

We expected the new words to elicit a frontal N400 reflecting the engagement of prefrontal regions associated with working memory (Hagoort, 2019) and meaning acquisition (Rodríguez-Fornells et al., 2009). Besides, we expected a decrease of the N400 throughout the learning, reflecting the gradual build-up of associative memory traces. As novel words were increasingly consolidated into the mental lexicon (Dumay & Gaskell, 2012), we also expected a shift of the topographical distribution of the N400 from frontal to parietal regions reflecting the consolidation of the words.

In *Experiment 2*, we replicated the findings of *Experiment 1* and ruled out the effects of stimulus repetition and familiarity on the N400. We used a between-group design with 38 new participants randomly assigned to a Cons group or a Incons in a single training session. Participants of the Cons were presented with consistent object-word associations, whereas the Incons group of participants were presented with inconsistent associations rendering learning impossible. We hypothesized the N400 would be modulated throughout the training session in the Cons but not in the Incons, thus, reflecting the integration of the novel word-form rather than a repetition/familiarity effect (Bermúdez-Margaretto, 2015).

2.2 Study II: Oscillatory activity and functional connectivity of concurrent word segmentation and meaning mapping in 9-year-old children

In study II, we aimed to explore the functional role of brain oscillations and functional connectivity during concurrent speech segmentation and meaning mapping in 9-year-old children. Classically, speech segmentation and meaning acquisition have been considered as two independent steps of language learning. Interestingly, these two processes occur in parallel in adults (François et al., 2017). Moreover, it is well known that multisensory information such as visual cues can facilitate speech segmentation (Cunillera et al. 2010b, Thiessen et al., 2010) and meaning mapping (Tomasello, 2000; Gleitman, 1990). Previous studies suggest that brain oscillations play a crucial role in various cognitive functions, including stimuli processing

Research aims

(Canolty et al., 2006; Lakatos et al., 2005). Here, we aimed to explore the neural correlates of simultaneous speech segmentation and meaning acquisition in children. Sixty 9-year-old children performed an audio-visual SL task while we recorded EEG. Children were exposed to (i) a structured stream with consistent word-picture associations (Cons condition) and (ii) a random stream with inconsistent word-picture associations (Incons condition). After the training, participants were tested on their ability to recall words and word-picture associations. We used frequency tagging and connectivity analyses based on wPLI measures to explore modulations of neural synchronization and long-distance connectivity during learning. We expected children could simultaneously segment words and associate them with meaning. We hypothesized that the simultaneous presentation of consistent picture-words association should enhance this process. Moreover, based on previous studies showing that an increased neural synchronization is related to better comprehension and speech processing (Luo & Poeppel, 2007; Peele et al., 2013), we expected an increase of neural oscillatory activity at word and syllable rate reflecting the processing and integration of the words. Importantly, as the word onset occurred simultaneously to the picture presentation, we hypothesized that neural synchronization with the stimuli should be facilitated by visual cues that direct learners' attention to the words onsets.

2.3 Study III: The role of the hippocampus in SL

The third study was carried out in the Hospital “La Timone” (Marseille, France) during the international Ph.D. internship I made under the supervision of Dr. Daniele Schön, team leader of the Dynamic of Cognitive process from the Institute of System Neuroscience. The project's goal was to explore the role of the MTL and especially of the hippocampus for speech segmentation based on SL. Although controversial, several studies have suggested that the hippocampus could be involved in SL by playing an important role in extracting pattern regularities in the visual modality (Turk-Browne et al., 2009, Schapiro et al., 2014; Henin et al., 2020). However, the involvement of the hippocampus in auditory SL is less clear (but see Schapiro et al., 2014).

Here, we collected intracranial recordings in seven pharmaco-resistant temporal lobe epilepsy patients to directly assess the role of the hippocampus during speech segmentation based on SL. Patients were exposed to a stream of pseudowords with the TP between words as the only cue for word segmentation while we recorded stereotactic EEG. We used frequency tagging analysis to test the level of neural synchronization in auditory and hippocampal regions at syllables and word rates. Indeed, better neural alignment with external stimuli has been

Research aims

associated with better stimulus processing and speech comprehension (Peele et al., 2013; Assaneo & Poeppel, 2018; Ahissar et al., 2001; Luo & Poeppel, 2007; Nozaradan et al., 2011;2014). We hypothesized that cortical auditory regions would show high neural synchronization to syllable frequency, whereas the hippocampus would respond to word frequency instead.

Chapter 3

Study 1

This study corresponds to:

Neus Ramos-Escobar, Matti Laine, Mariana Sanseverino-Dillenburg, David Cucurell, Clément François, Antoni Rodríguez-Fornells. The interplay between domain-general and domain-specific mechanisms during novel word learning: An Event-Related Potential Study

Chapter 3 Study 1

The interplay between domain-general and domain-specific mechanisms during novel word learning: an event-related potential study

Humans continuously learn new information. Here, we examined the brain dynamics of explicit novel word learning. In a first experiment, 25 adults learned word-object associations during a 5-day training program allowing us to track the N400 dynamics across learning blocks within and across days. Successful learning was accompanied by an initial frontal N400 that decreased in amplitude across blocks during the first day and shifted to parietal sites during the last training day. In *Experiment 2*, we replicated our findings with 38 new participants randomly assigned to a consistent learning (Cons) or an inconsistent learning group (Incons). Our results show that N400 amplitude modulations, both within and between learning sessions, signal the emergence of novel lexical traces in the absence of facilitation from pre-existing semantic knowledge. The shift in N400 topography suggests that different N400 neural generators may contribute to specific word learning steps by relying on a balance between domain-general and language-specific mechanisms.

Keywords: word-learning, memory, EEG, N400, vocabulary acquisition

Study 1

3.1 Introduction

Humans present an exceptional ability to learn. In the case of language, this is evident when looking at the capacity to learn novel words. However, compared to infants' first language acquisition, word learning in adulthood appears as a non-automatic, demanding task that requires motivation and active control (Abutalebi, 2008). Recent cognitive learning theories describe a gradual shift from highly controlled to more automatized processing during the acquisition of new knowledge or skills (Chein & Schneider, 2005; 2012). Interestingly, the early stages of word learning are also known to require not only the coordinated engagement of specific language and memory systems but also domain-general regulatory functions such as attention, cognitive control, and motivation (Laine & Salmelin, 2010; Rodríguez-Fornells et al., 2009; Sliwiska et al., 2017). These domain-general mechanisms can regulate and monitor specialized cortical networks involved in word learning (Hagoort, 2019; Abutalebi et al., 2013). Thus, the subtle balance between domain-general and domain-specific neural resources might provide a highly flexible system involving both short- and long-term brain plasticity to integrate and automatically retrieve the learned information (Chein & Schneider, 2012; Jeon & Friederici, 2015).

Psycholinguistic models propose that word learning may take place at different time scales, from the rapid learning of object-word associations to the slow build-up and consolidation of new episodic memory traces (Davis & Gaskell, 2009; McClelland et al., 1995; Sharon et al., 2011; Coutanche & Thompson-Schill, 2014). Sleep-related consolidation mechanisms are crucial for the binding of new words onto pre-existing lexical-semantic information (McClelland et al., 1995; Davis & Gaskell, 2009; Havas et al., 2018; Sestieri et al., 2017; Tamminen & Gaskell, 2013). Importantly, recent evidence suggests that rapid cortical encoding of new information is independent of the middle temporal lobe (Hebscher et al., 2019; Brodt et al., 2018).

Capitalizing on its remarkable capacity to track the temporal brain dynamics of cognitive computations, event-related brain potentials (ERPs) have been extensively used to decipher word-learning mechanisms (McLaughlin et al., 2004; Dittinger et al., 2016; Mestres-Missé et al., 2007; Stein et al., 2006; Yum et al., 2014; Ferreira & Bernaldes, 2020; Morgan-Short, 2014). One of the most reliable findings observed in adults, children, and infants, is the modulation of the N400 component during learning new words (Friedrich & Friederici, 2004; 2017; Torkildsen et al., 2008; McLaughlin et al., 2004; Dittinger et al., 2016; Mestres-Missé et al., 2007; Stein et al., 2006; Elgort et al., 2015; Yum et al., 2014; Kuipers et al., 2017; Perfetti, 2007; Soskey, et al.,

Study 1

2016), a component associated with lexical and semantic processing (Kutas & Federmeier, 2011). However, discrepancies in the topographical distribution of the N400 and interpretations of its functional role persist. For instance, new word learning is associated with N400 modulations over prefrontal regions, reflecting the engagement of prefrontal cognitive control mechanisms for monitoring the initial build-up of episodic memory traces for novel words (Stein et al., 2006; Yum et al., 2014). Because newly created traces and associations are still weak, general cognitive control mechanisms may support the maintenance of novel information in short-term memory during effortful retrieval. In contrast, at more advanced learning stages following the initial memory consolidation, processing of newly learned words may become more automatized, less susceptible to interference, and less dependent on cognitive control and monitoring processes during encoding and retrieval. In line with these ideas, several fMRI studies have shown that during the initial stages of learning new words, activation is observed in cognitive control regions (Bradley et al., 2013; Mestres-Missé et al., 2008; 2009; Paulesu et al., 2009; Clements-Stephens et al., 2012; Ripollés et al., 2014; Ferreira et al., 2015; Yang et al., 2015; Hosoda et al., 2013; Tagarelli et al., 2019). In a similar vein, it has been proposed that early bilinguals may compensate for their lack of proficiency by an increase in cognitive control, probably needed for the internal management of conflicts between competing for lexical representations (Briellmann et al., 2004; Pillai et al., 2003; Abutalebi & Green, 2007; Rodríguez-Fornells et al., 2009, Li et al., 2014).

Here, we designed a longitudinal study that allowed us to track the N400 modulations from the beginning of word-referent learning until the later stages of learning after consolidation has taken place (after five days and four months post-training). Based on previous studies, we expected the new words to elicit a frontal N400 during day 1. We expected the N400 amplitude to decrease through learning blocks, reflecting the gradual build-up of associative memory traces. Moreover, as overnight consolidation should integrate the new words into the mental lexicon (Havas et al., 2018; Dumay & Gaskell, 2012) that relies upon language-related cortical networks, we expected that improved recall performance through learning sessions would be accompanied with an N400 topography shift from frontal to parietal regions, reflecting an increase in retrieval automaticity. An innovative aspect of this study was that participants had to associate new words to unknown objects, representing a particularly demanding form of novel word learning. Therefore, ERP modulations observed in the present study cannot rely on the activation of pre-existing lexical memory traces (Angwin et al., 2014). Finally, we explored

Study 1

individual differences in cognitive control that may explain our behavioural and ERP data (Linck & Weiss, 2015; Kapa & Colombo, 2014).

In addition, we conducted a second ERP experiment to rule out the confounding effects of stimulus repetition and familiarity on the N400 evolvment. We randomly assigned new participants to two groups differing in the consistency of the to-be-learned word-referent associations. While participants of the Cons group were presented with consistent word-object associations, participants of the Incons group were presented with inconsistent associations rendering learning impossible. If the N400 amplitude decrease observed in *Experiment 1* does not result from a stimulus repetition/familiarity confound, a decrease of the amplitude of the N400 amplitude through learning blocks should be seen only in the Cons group.

3.2 Materials and methods

3.2.1 Participants

Twenty-five healthy volunteers (13 females, mean age: 22 ± 3.19) participated in *Experiment 1*. They were all right-handed and had normal or corrected-to-normal vision, and reported no history of neurological deficits. All participants were Spanish-Catalan bilinguals except for two Spanish monolinguals. For the follow-up evaluation, 5 participants dropped out of the study (N=20).

For *Experiment 2*, a total of 38 healthy volunteers were recruited with 19 participants randomly assigned to the learning group (Cons, 10 females, age: 23 ± 3.32) and the other 19 participants to the no learning group (Incons, 9 females, age: 23 ± 2.49). They were all right-handed, had normal or corrected-to-normal vision, and reported no history of neurological deficits. All participants were Spanish-Catalan bilinguals.

For both experiments, participants were informed about the study procedure, gave their written consent, and were paid for their participation. The study was approved by the local ethics committee.

3.2.2 Stimuli

A set of 139 unfamiliar black-and-white objects representing farming artifacts were selected from the Ancient Farming Equipment (AFE) word-learning paradigm (Laine & Salmelin, 2010).

Study 1

Familiarity scores for these 139 unfamiliar objects (0-5 from totally unfamiliar to very familiar) were obtained from 20 additional students from the University of Barcelona (Mean: 2.58; SD: 0.61). Based on these ratings, the 120 most unfamiliar objects were selected.

A set of 120 tri-syllabic pseudo-words was created using the B-Pal software (Davis & Perea, 2005). The pseudo-words respected the phonotactic rules of Spanish, representing six different Consonant-Vowel (CV) structures (CVCVCV, VCVCVC, VCVCV, VCCVCV, CVCVCCV, and VCVCCV).

3.2.3 Procedure

Experiment 1

Experiment 1 included five consecutive daily training sessions and a follow-up evaluation four months later (see **Figure 23**). During the first and last days of the training program, the participants first performed a pre-exposure task during which they were exposed to novel objects and pseudo-words, (A) a learning phase with (B) an overt-naming task in-between the learning. Consecutively, the participants performed a covert-naming task (C), an overt-naming task (B), and a 2-alternative forced-choice word-object matching task (2-AFC, D1; see below for detailed task description). During the second, third, and fourth training days, the participants performed the learning task followed by the 2-AFC task. In the 4-month follow-up assessment, the participants performed the overt-naming and the 2-AFC tasks. EEG was recorded during Day 1 and Day 5.

On the first day, before starting the training sessions, we administered a Language history questionnaire (Bilingualism and Language switching, Rodríguez-Fornells et al., 2012) and two tasks associated with cognitive control and working memory [the semantic and phonological fluency tasks (category of animals and words beginning with the letter "p") and an auditory working memory task (WAIS-III's Digit Span, mean score of forward and backward span). Verbal fluency was chosen because of its previous association with both aspects, effortful word retrieval and cognitive control (Hughes & Bryan, 2002; Carpenter et al., 2020). More importantly, verbal fluency has been consistently associated with the amplitude of the N400 component during lexical retrieval (Federmeier et al., 2002, 2010). Digit span was chosen as a measure of working memory and cognitive flexibility, which has previously related to word learning (Baddeley, Gathercole & Papagno, 1998). After this brief examination, participants underwent the training sessions.

Study 1

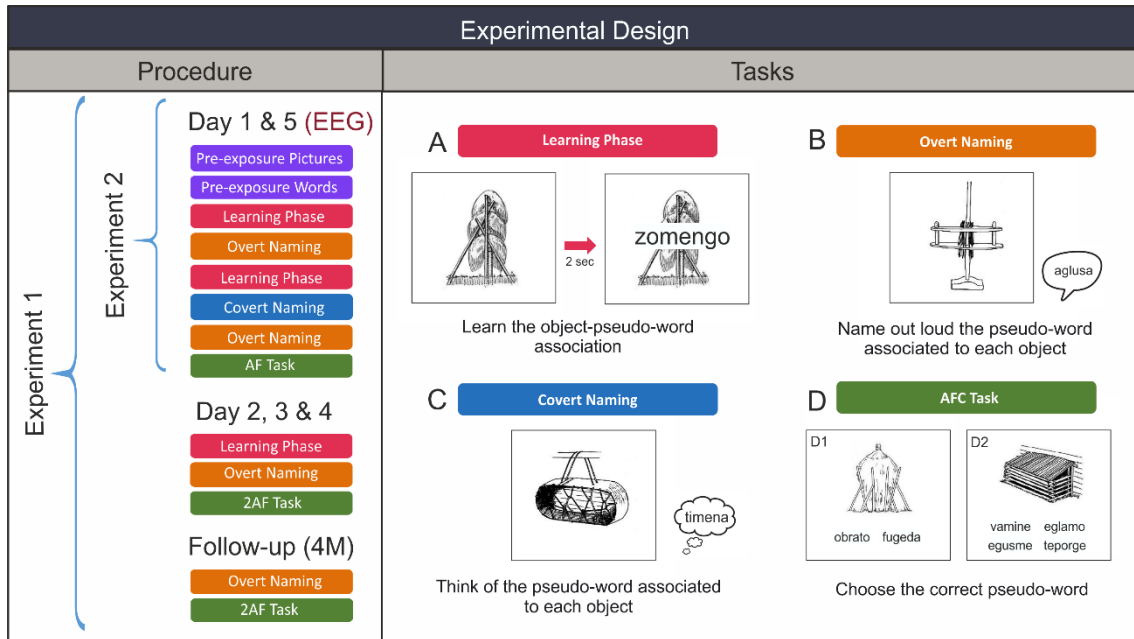


Figure 23. Overview of the experimental design. In *Experiment 1*, participants performed a series of tasks (A-D) during a 5-day training program and a 4-month follow-up. In *Experiment 2*, the procedure was similar to Day 1 of *Experiment 1*. (A) Learning phase, (B) Overt-naming task, (C) covert-naming task, (D) two alternative-forced choice task (2-AFC). In *Experiment 2*, participants performed the same series of tasks from A to D. Note that a 4-AFC task was used in *Experiment 2*.

First, in the pre-exposure phase, the 120 novel objects were randomly presented for 2 seconds each, together with 5 familiar objects from the AFE paradigm (objects rated as familiars in the initial item selection). The participants were asked to press a response button when they saw a familiar object to maintain a constant level of attention throughout the pre-exposure phase. This task aimed to familiarize participants with the novel words and pictures before and after learning the associations. A similar task was performed for the pseudo-words. The 120 novel words were randomly presented for 2 seconds each, together with 5 real words, and the participants were asked to press a response button when a real word appeared. The results of this task are not presented in the present work.

In the learning phase, the 120 object-word pairs were randomly presented within each block. The participants were required to pay attention to the object-word associations and encouraged them to learn as many as they could. In each learning trial, a novel object was presented on the screen for 2 seconds, and the associated pseudo-word appeared on the top of it for 1 second. The first part of the learning phase consisted of 480 trials divided into 4 blocks, while the second part (after the overt naming task) consisted of 240 trials divided into 2 blocks. During the learning phase, a short break was taken every 30 trials to avoid fatigue. The overt-naming task was done shortly before the end of the learning task to boost the learning.

Study 1

The level of learning was assessed using three different tasks.

Overt-naming task: each of the 120 objects was presented on the screen for 3 seconds, while the participants had to name out loud the pseudo-word associated with the object. To control for possible order effects, eight randomized sequences of the objects were created, and the presentation of these sequences was counterbalanced across the different days and participants. The experimenter was seated next to the participant monitoring their naming responses. We considered the response as correct when the participant named the object without any spelling mistake. The EEG data obtained in this task are not presented in this paper.

Covert-naming task: each of the 120 objects were presented on the screen for 3 seconds, while the participants were asked to think about the pseudo-word without overtly naming it. To maintain the level of attention constant during the task, a red square appeared around the object in 10% of the trials, prompting the participants to name that object overtly. For this task, 5 randomized versions were created (one version for each time that the task was administered) and counterbalanced across sessions and participants. The EEG data obtained in this task are not presented here.

2-AFC task: each of the 120 objects was presented on the screen for 2 seconds, together with two possible pseudo-words appearing below, and followed by a question mark for 500 ms prompting the participants to respond. The participants had to choose which of the two pseudo-words was the one associated with the object by pressing the right or the left mouse button. One of the two pseudo-words was the correct choice, and the other one had been presented during the learning phase but had never been associated with that object. Two randomized sequences were created and counterbalanced across sessions and participants.

Experiment 2

In *Experiment 2*, the procedure was identical to the one used in Day 1 of *Experiment 1*, except for small changes in the learning and the 2-AFC task. In the learning phase, the Cons group was exposed to fully consistent object-word associations as in *Experiment 1*. By contrast, the Incons group was presented with fully inconsistent associations so that none of the 120 words was presented more than once with the same object (see **Figure 23**). Therefore, no learning was possible. Both groups were instructed to learn as many associations as they could. The task was the same as in *Experiment 1* except that they also had to press a button each time one of six previously familiarized objects appeared (catch trials). The catch trials appeared twice in each

Study 1

block and were pseudo-randomly divided in the sequence with a minimum of two learning trials between the familiarized objects. Those six additional novel objects were selected from the AFE paradigm (catch trials). This extra task was done to maintain the participants' level of attention constant throughout the learning phase. Before the beginning of the experiment, the participants received these 6 objects by e-mail and were asked to memorize them. At the beginning of the session, we confirmed sufficient familiarization with these stimuli by presenting the six objects together with fourteen extra objects (which were not used again in the experiment), and participants were required to recognize them. The duration of the trials and the block structure were the same as in *Experiment 1* for both groups.

For the AFC task, we used a four-alternative word-object matching task (4-AFC) instead of 2-AFC to avoid the ceiling effect obtained in *Experiment 1*. The 120 objects were randomly presented for 3 seconds together with 4 possible pseudo-words below, followed by a question mark for 500 ms prompting the participants to respond. The participants were asked to press buttons 1 to 4 to indicate which word was associated with the new object. For the Cons group, only one of the pseudo-words was correct, while the 3 other pseudo-words were presented during learning but never with that specific object. For the Incons group, one of the six pseudo-words had been associated with the object during the learning task, and the other three pseudo-words had never been associated with that object. The correct pseudo-word in the Incons group was counterbalanced across blocks and tasks so that all pseudo-words were presented the same number of times.

The pre-exposure, overt-naming, and covert-naming tasks were identical to the ones in *Experiment 1*. The EEG was recorded during the entire experiment. The pre-exposure task was maintained in such a way that the number of pictures and word repetitions was similar in both experiments.

3.2.4 EEG data acquisition and analyses

The EEG signal was recorded from the scalp using tin electrodes mounted in an electrocap (Electro-Cap International) and located in 29 standard positions during Day 1 and Day 5 in *Experiment 1* and during the entire session in *Experiment 2*. Biosignals were re-referenced off-line to the left and right mastoidal electrodes. Vertical eye movements were monitored with an electrode placed at the infraorbital ridge of the right eye. Electrode impedances were kept below 5 k Ω . The electrophysiological signals were digitalized at a rate of 250 Hz.

Study 1

Electrophysiological data were analyzed using ERPLAB 13.5.4b. The EEG was filtered off-line using a 30 Hz low-pass filter only for display figures. Epoch rejection criteria were individually determined using a simple voltage threshold within a range of +/- 50 for eye electrode and +/- 75 μ V for the other channels and forward visually checked for each trial and participant. For the analysis, we focused on the learning phases of the experiment. Epochs of 900 ms were time-locked to both object and word presentation considering a -100 ms pre-stimulus baseline.

3.2.5 Behavioral data analysis

For *Experiment 1*, we analyzed separately for each task (overt-naming & 2AFC task) the percentage of correct responses (CR) using repeated-measures ANOVA with Day (Day 1, 2, 3, 4, 5, and 4-months follow-up) as a within-subject factor. The Greenhouse-Geisser correction was applied when necessary. Furthermore, as a measure of new-word Forgetting, we computed the words retrieved on Day 5 minus the words retrieved in the 4-month follow-up evaluation divided by the number of words retrieved on Day 5 (Day5 - Follow-up / Day5). This index was used to explore the relationship between the participant's ability to learn the words and neuropsychological factors. Only 20 participants were included due to the drop of 5 participants in the follow-up evaluation. Pearson correlations were calculated using IBM SPSS Statistics 23 software (SPSS, Chicago, Illinois).

For *Experiment 2*, we analyzed separately for each task (overt-naming & 4AFC task) the percentage of correct responses using independent-samples t-tests. We also used one-sample t-tests for each group to determine whether participants' performance was significantly above chance level. For the Icons group, we considered a correct answer in the overt-naming task when a participant named at least one of the six different pseudo-words presented during the learning blocks. All the analyses were performed with IBM SPSS statistics 23 (SPSS, Chicago, Illinois).

3.2.6 ERP analyses

In both experiments, the four blocks of the first part of the learning task were averaged in 2 blocks (Block 1 and Block 2), and the two blocks of the second part (after the overt naming) were averaged in 1 block (Block 3). In *Experiment 1*, based on visual inspection of the waveforms as well as on previous literature (Soskey et al., 2016; see Kutas & Federmeier, 2011 for a review), we analyzed the N400 elicited by words in the 350-550 ms time-window. Only correctly learned object-word associations (correct items in the 2-AFC task of Day 1) were included in the analyses.

Study 1

We performed repeated-measures ANOVA on nine channels (F3, C3, P3, Fz, Cz, Pz, F4, C4, and P4) with four factors: Day (Day 1, Day 5), Block (Block 1, Block 2, Block 3), Lateralization (Left, Central, and right) and Antero-posterior (Frontal, Central, and Parietal).

To investigate the possible association between the main ERP effects and behavioral and cognitive control individual differences (verbal fluency and working memory), we computed the mean amplitude of the N400 difference waveform of the within (Block 1 minus Block 3 of Day 1) and between-session effect (Block 1 minus average of Day 5) over a frontal region of interest. For both effects, we took the electrodes where the topographical distribution showed larger effects. For the within-session effect, with a central parietal distribution, we included Pz, P4 & Cp2 (see **Figure 26A**), and for the between-sessions effect, with a frontal distribution, we took Fz, Fcz, Cz, Fc1 & Fc2 (see voltage map of **Figure 26B**). Pearson correlation was used for the analysis, and False Discovery Rate (FDR) correction was applied for each neuropsychological test and each behavioral measure.

Finally, we computed current source density (CSD) analysis to localize the patterns of the current flow of the N400 using the CSD toolbox (Kayser & Tenke, 2006) in Matlab.

In *Experiment 2*, we included all the trials in the analysis due to the mixed-model design. We considered the same time window as in *Experiment 1* for the N400. An ANOVA was used with Group (Cons vs. Icons) as between-subject factor and Block (Block 1, Block 2, and Block 3), Lateralization (Left, Central, and Right) and Antero-posterior (Frontal, Central, and Parietal) as within-subject factors including the same electrodes as in *Experiment 1*.

For the Cons group only, we analyzed ERP differences in N400 amplitude between learned and non-learned words by averaging the ERPs throughout blocks and by performing a repeated-measures ANOVA with the factors Item (learned words and non-learned words), Lateralization (Left, Central, and Right) and Antero-posterior (Frontal, Central, and Parietal). For each ANOVA, Greenhouse-Geisser sphericity corrections were applied when appropriate. Post-hoc tests were conducted using pairwise t-tests corrected for multiple comparisons using the Benjamini-Hochberg FDR procedure.

Study 1

3.3 Results

3.3.1 Experiment 1

Behavioral data

For *Experiment 1*, we analyzed separately for each task (overt-naming & 2-AFC task), the percentage of correct responses (CR) using repeated-measures ANOVA with Day (Day 1, 2, 3, 4, 5 and 4-months follow-up) as a within-subject factor (**Figure 24**). Results showed that the participants increased their number of CR over the 5 days of training in both tasks: the overt-naming task [main effect of Day: $F(5, 95) = 145.89, p < .0001, r^2 = .85$] and in the 2-AFC task as well [main effect of Day: $F(5, 95) = 43.89, p < .0001, r^2 = .69$].

Remarkably, although all participants were able to learn along with the training sessions, the number of words that participants were able to retrieve in the follow-up evaluation was quite low (forgetting: mean = .93, SEM = .01, **Figure 24A**). Nonetheless, the average rate of correctly recognized words in the 2AFC task was quite high (mean = 77.66, SEM = 1.76; **Figure 24B**), which suggests that considerable information was retained in passive vocabulary. Interestingly, we found a significant correlation between the forgetting index and semantic fluency performance ($r = -0.47, p = .03$), so that a better performance in semantic fluency was associated with less forgetting in the follow-up evaluation (**Figure 24C**). No other significant correlations with the other neuropsychological tests (phonological fluency or working memory) were found.

Study 1

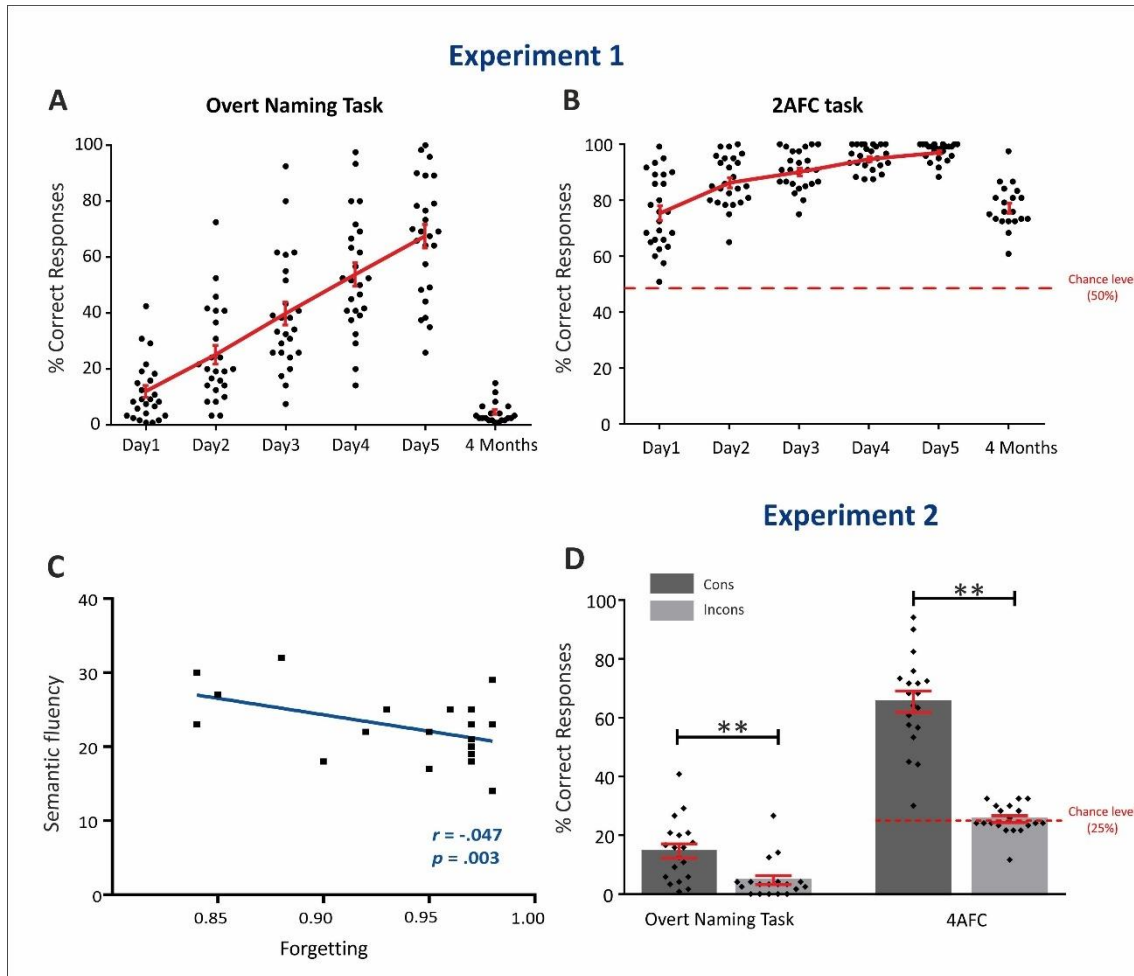


Figure 24. Behavioral data. Experiment 1: (A) Percentage of correct responses in the overt-naming task across sessions; (B) Percentage of correct responses in the 2-AFC task; (C) The negative correlation between the score in semantic fluency and 4-month forgetting. (D) Experiment 2: Percentage of correct responses in both the overt-naming task and the 4-AFC task for the consistent learning (Cons) and the inconsistent learning group (Incons). Dots and diamonds represent individual values and bars correspond to the mean and standard error of the mean (SEM). Notice that in section C, only 20 participants were included in the correlation due to the drop off of 5 participants in the follow-up evaluation.

ERP results

Results of the repeated measures ANOVA with the factors Day (Day 1, Day 5), Block (Block 1, Block 2, Block 3), Lateralization (Left, Central, and right) and Antero-posterior (Frontal, Central, and Parietal) are shown in **Table 1**. A frontally distributed N400 was elicited by the presentation of new words during the first block of Day 1 (see **Figure 25**). This frontal N400 was modulated by learning with a decrease of amplitude in Block 2 and 3, as evidenced by a main effect of Block [$F(2,48) = 14.39$, $p < .001$] and a significant Day x Block interaction [$F(2, 48) = 14.49$, $p < .001$]. The decrease in N400 amplitude across blocks was significant in Day 1 [Day 1: Block 1 vs Block 2 ($t(24) = -3.05$; $p = .006$); Block 2 vs Block 3 ($t(24) = -4.15$, $p < .001$)] but not in Day 5 during which

Study 1

it reached a plateau [Day 5: Block 1 vs Block 2 ($t(24) = 1.89, p = .07$); Block 2 vs Block 3 ($t(24) = -1.56, p = .132$)]. This was further evidenced by separated analyses by day showing a significant effect of Block during day 1 only [Day 1: main effect of Block $F(2, 48) = 23.17, p < .001$; Day 5: main effect of Block: $k(2, 48) = 2.15, p = .13$].

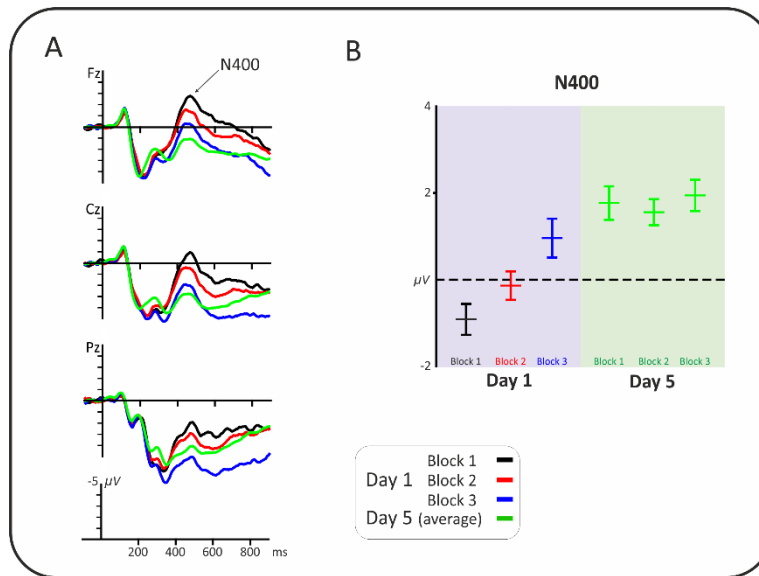


Figure 25. ERP grand averages to new words at midline electrodes across 25 participants during each learning block (black = block 1 of Day 1, red = block 2 of Day 1, blue = block 3 of Day 1, green = average of the three blocks of Day 5). B) Time-course of within- and between-session N400 mean amplitude for all the learning blocks of Day 1 and Day 5 in the 350-550 ms time-window and over the frontal regions (averaged over F3, Fz, & F4 electrodes).

Importantly, the scalp distribution of the N400 during the first block was frontocentral (see **Figure 26A**, Block 1 Day 1 topography in the right panel) and showed a shift in the distribution along with the experiment. The within-session N400 effect (difference waveform of block 1 minus block 3 of Day 1) showed a centro-parietal displacement (**Figure 26A**, voltage maps & CSD panels). In contrast, the between-session N400 effect (difference waveform of Block 1 of Day 1 minus the average of Day 5) showed a very clear medial frontal topography [Day x Block x Antero-posterior interaction: $F(4,96) = 4.78; p = .013$; see **Table 2** and **Figure 26B**, voltage maps & CSD panels].

Regarding the relationship between individual differences and ERP effects, we observed a significant negative association between the N400 frontal effects (difference waveform for the between-session effect; see Fig. 4B) and phonemic fluency performance ($r = -.45, p = .025$, **Figure 27**), so that participants with higher phonemic fluency performance showed greater changes in the N400 between-sessions (frontal distribution). No significant correlations were found

Study 1

between the other cognitive tasks or learning performance with the within-session N4 effect (parietal distribution).

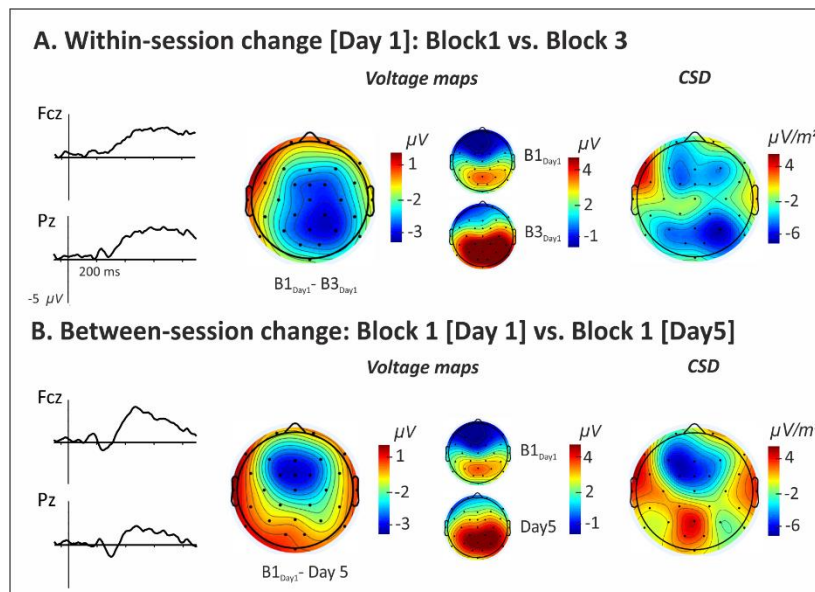


Figure 26. Difference waveform, topographical distribution, and current source density maps (CSD) of the N400 within- and between-sessions between 350-550 ms. A) Distribution of the difference waveform of the within-session N400 effect (Block 1 - Block 3 of Day 1). B) Distribution of the difference waveforms between-sessions N400 effect (Block 1 of Day 1 - the average of all Blocks of Day 5). For each subfigure, the difference waveform in Fcz and Pz electrodes is shown on the left. The voltage maps for the single conditions used to compute the difference waveform are shown in the middle. Notice the waveform change from a central-parietal N400 within-session distribution to a more frontally distributed FN400 between-sessions. The CSD maps of the difference waveforms are shown on the right.

Regarding the relationship between individual differences and ERP effects, we observed a significant negative association between the N400 frontal effects (difference waveform for the between-session effect; see Fig. 4B) and phonemic fluency performance ($r = -.45$, $p = .025$, **Figure 27**), so that participants with higher phonemic fluency performance showed greater changes in the N400 between-sessions (frontal distribution). No significant correlations were found between the other cognitive tasks or learning performance with the within-session N400 effect (parietal distribution).

Study 1

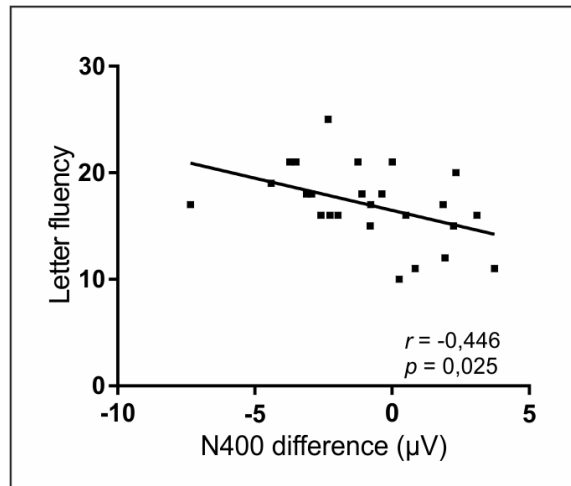


Figure 27. Correlation between the between-session N400 amplitude modulation (Block 1 of Day 1 minus average of Day 5) over a frontal region of interest (Fz, Cz, Fcz, Fc1, Fc2) and participant's level of performance in the verbal fluency letter task.

3.3.2 Experiment 2

Behavioral data

For *Experiment 2*, the Cons outperformed the Icons group in both tasks (4-AFC task: Cons: 65.44% CR, SD = 15.7; Icons: 25.48% CR, SD = 5.04; overt-naming task: Cons: 14.65%, SD = 10.44; Icons: 4.74%, SD = 6.54; see **Figure 24D**) as evidenced by significant group differences in the overt-naming [$t(36) = 3.51, p = .001$] and in the 4-AFC task [$t(36) = 10.52, p < .001$]. These results were confirmed by further pairwise t-tests showing significant group differences in both the overt-naming [$t(36) = 4.93; p = .001$] and the 4-AFC task [$t(36) = 11.93; p < .001$]. Importantly, in the 4-AFC, the Cons group (M = 65.44, SD = 15.76) performed significantly above chance level (25% of correct responses; $t(18) = 11.18, p < .001$) which was not the case for the Icons group (M = 25.48, SD = 5.04; $t(18) = .42, p < .681$), indicating that the latter group was not able to associate the words to the objects.

ERP results

Results of the repeated-measures ANOVA with the factors Item (learned words and non-learned words), Lateralization (Left, Central, and Right), and Antero-posterior (Frontal, Central, and Parietal) are shown in **Table 3**. We observed different learning dynamics between the two groups at the level of the N400 [Group x Block interaction: $F(2,72) = 3.98, p = .02$]. Separate analyses in the Cons group showed differences across blocks [Block x Antero-posterior interaction: $F(4,72) = 6.28, p = .002$], thus confirming the results obtained in *Experiment 1*. New

Study 1

words elicited a frontal N400 in Block 1 that decreased in Block 2 and 3 (**Figure 28A**), with a significant decrease in Block 3 ($t(18)=-3.88$; $p = .002$, corrected for multiple comparisons). As found in *Experiment 1*, the within-session N400 effect was centro-parietal (**Figure 28C**). In the Icons group, the main effect of Block was significant [$F(2,36) = 3.95$, $p = .03$], but pairwise t-tests failed to reach significance after correction for multiple comparisons (see **Figure 28A**).

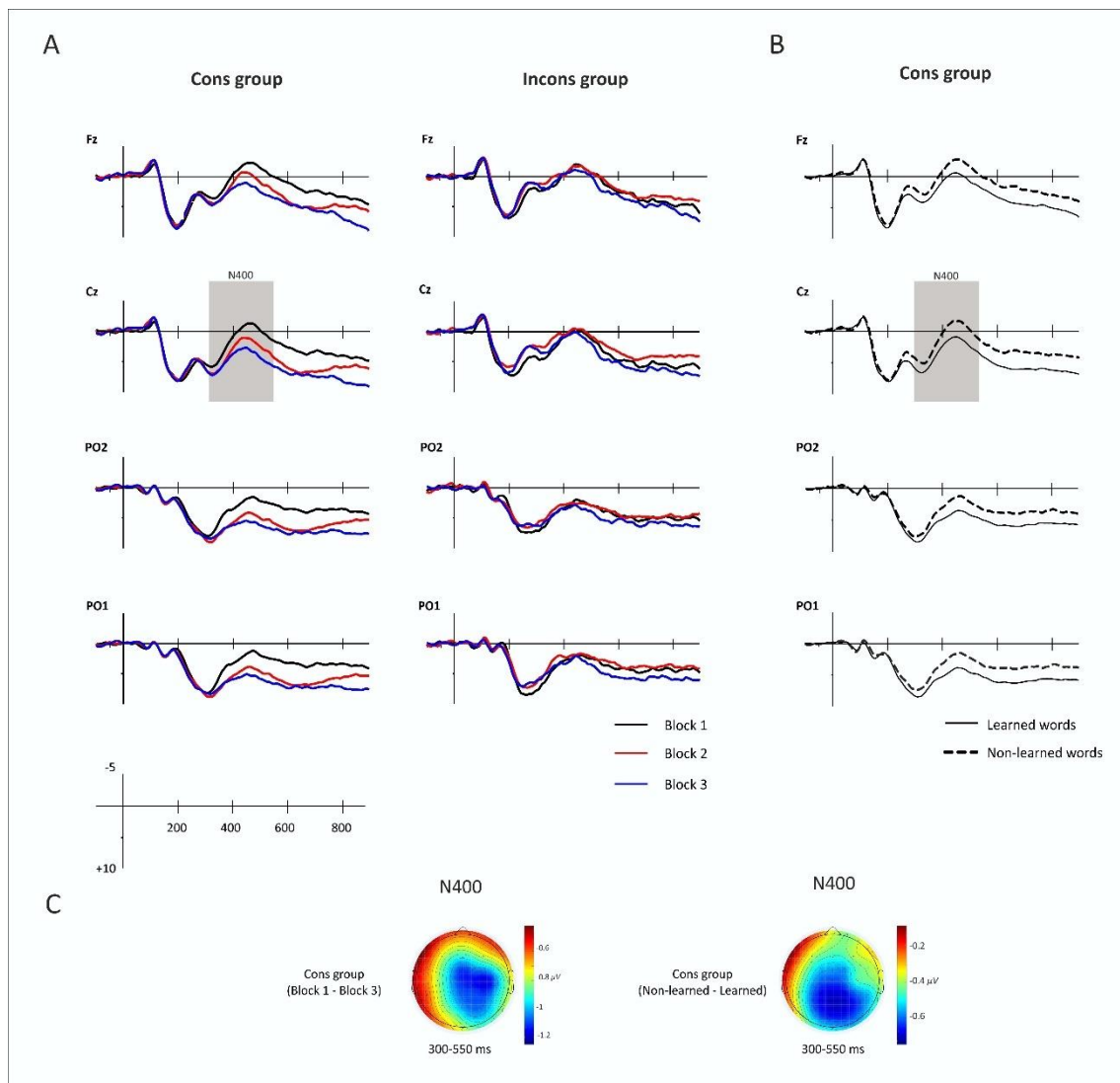


Figure 28. Grand average ERPs to new words in both the Cons (left) and the Icons group (right) at midline electrodes for each learning block. Notice that the N400 amplitude decreases across blocks in Cons but not in Icons. B) Grand average ERPs to words averaged across all blocks comparing correctly learned to non-learned words in the Cons group over midline electrodes. C) Scalp distribution of the N400 for the difference (Block 1 - Block 3) in the Cons group and the difference (Correct - Incorrect).

We further investigated the functional role of the N400 in the build-up of new semantic associations in the Cons group by comparing ERPs to correctly learned and non-learned words averaged across blocks. Results of the ANOVA revealed that non-learned words elicited a larger

Study 1

N400 than learned words between 350 and 550 ms [main effect of Item: $F(1,18) = 36.82$; $p < .001$, see Figure 28B]. We also found a significant Item x Lateralization interaction [$F(2,36) = 4.87$; $p = .021$]. The N400 effect (Incorrect – Correct) was maximal over central electrodes ($-2.1 \mu V$), which was confirmed by pairwise t-test showing a significantly larger effect over central than over left electrodes [$t(18) = 4.33$; $p = .001$, corrected for multiple comparisons]. However, no significant differences were found when comparing central to right electrodes [$t(18) = 1.74$; $p = .13$, corrected for multiple comparisons] and left to right electrodes [$t(18) = -0.97$; $p = .34$, corrected for multiple comparisons]. The N400 distribution for the difference waveform of non-learned and learned items was centro-parietal (Figure 28C).

3.4 Discussion

This study examined the time course of early encoding and long-term consolidation of new lexical information in human adults. We tracked novel word learning and concomitant brain modulations from the early stages of learning up to the later stages after consolidation had taken place. We also found significant relationships between word learning (performance and ERP effects) and verbal fluency measures that tap cognitive control during lexical retrieval. Our study provides converging evidence for fast and slow neuroplasticity mechanisms supporting the learning of new information that may reflect the dynamic interplay between language-specific and domain-general brain resources.

Electrophysiological evidence for fast plastic changes was provided in the first experiment and replicated in the second one. In *Experiment 1*, novel words elicited a frontal N400 during the first block of the learning program (see **Figure 25A**; **Figure 26A**, **Figure 26B**, right panel; & **Figure 28A** left panel). This result confirms previous studies in adults showing larger frontal N400 for learned than for non-learned or familiar words (McLaughlin et al., 2004; Perfetti et al., 2005; Frishkoff et al., 2010; Mestres-Missé et al., 2007; Batterink et al., 2011; Yum et al., 2014; Dittinger et al., 2016; Stein et al., 2006; Bakker et al., 2015). Importantly, when comparing the first and last learning blocks of Day 1 (see **Figure 26A**), we observed a within-session shift in the topographical distribution from frontal to centro-parietal regions. Dittinger and colleagues (2016) reported a different pattern of N400 topographical modulations during word learning, with an increase of N400 amplitude from the first to the second learning block within-session over left and midline electrodes. Differences in the topographical distribution between the two studies may arise from differences in the task and stimuli used. Alternatively, this apparent discrepancy may reflect the fact that frontal regions were largely engaged during the initial

Study 1

stages of learning (Day 1). Indeed, in our study, no stored semantic information was available for the new referents. Therefore, frontal activations related to effortful semantic processing and retrieval were likely to occur during the early stage of training (Petersen et al., 1988; Demb et al., 1995; Ferstl et al., 2008; Badre & Wagner, 2002; Gold et al., 2006; Thompson-Schill et al., 1997).

We also found evidence for slow neuroplastic changes associated with automatization and long-term consolidation of new associations in *Experiment 1*. Specifically, we observed that the between-session learning effect (i.e., the difference in N400 amplitude between the first learning block of Day 1 and the average of all the learning blocks during Day 5, see **Figure 25A** & **Figure 25B**) had a frontal distribution (see **Figure 26B**). Importantly, the N400 amplitude reached a plateau during Day 5, which may further suggest that the newly learned words had acquired robust and automatized lexical-semantic memory traces. Moreover, the increase in overt-naming performance, from low performance during the early stage of the training (11.5% in *Experiment 1* and 13.8% in *Experiment 2*) to quite high naming scores at the later stage of the training (74% during Day 5) supports this view. In the second experiment, we replicated the results of *Experiment 1* and found, again, a frontal N400 to novel words that decreased across learning blocks in the Cons group only (**Figure 28A**). Crucially, no modulations of the N400 across blocks were observed in the Icons group. Further, learned words elicited a smaller N400 than non-learned words despite equal exposure to both in the Cons group (**Figure 28B**). Although previous studies have shown that the N400 to non-words can be modulated by repetition (Bermúdez-Margaretto et al., 2018; Deacon et al., 2004), our results favor a lexical-semantic interpretation of the N400 modulation (Bermúdez-Margaretto et al., 2015, 2018, 2019; Kutas & Federmeier, 2000; Bakker et al., 2015). Our results confirm that semantic access was already taking place during the first day of training, probably due to a rapid learning process occurring during the initial acquisition of novel words (Rodríguez-Fornells et al. 2009).

Interestingly, we found an association between the frontal N400 component (block 1 of Day 1 minus average of Day 5) and the performance in a phonemic fluency task (see **Figure 27**). As the modulation of the N400 has been considered as a word learning index (McLaughlin et al., 2004; Mestres-Missé et al., 2007; Borovsky, Kutas & Elman, 2010), this result directly relates the proficiency in verbal fluency with the ability to learn novel words. This result is in line with the correlation found between semantic fluency and forgetting at 4 months (see **Figure 24C**). Results pointing out in this direction have been reported in recent studies (Qi et al., 2016; Carpenter et al., 2020). Some of these studies showed a relationship between working memory (Linck &

Study 1

Weiss, 2015) and inhibitory control processes with second language learning performance (Kapa & Colombo, 2014). Verbal fluency tasks require not only fast retrieval of words stored in the mental lexicon but also careful monitoring to select the appropriate words and avoid repetitions, i.e., cognitive control (Shao et al., 2014). Thus, the relationships between verbal fluency, ERP modulation during word learning, and forgetting found here converge with the idea of the involvement of cognitive control mechanisms during the early stages of word learning (Hill & Wagovich, 2020).

The initial steps of word learning are arduous, and hence cognitive control and top-down processes such as voluntary attention and motivation are more likely to contribute to learning during these stages (Rodríguez-Fornells et al., 2009; Laine & Salmelin, 2010). It has been shown that lexical processing during word reading is top-down modulated by the intention to engage in the task at hand (Strijkers et al., 2015). However, a gradual automatization occurs when learning advances (Chein & Schneider, 2012). Interestingly, for an associative word learning task critically, this automatization depends on sleep (Tham et al., 2015). Such an automatization may be reflected by increased resistance to semantic interference during a semantic judgment task (Kaczer et al., 2018) and by an improvement in proficiency known to depend on the activity of the prefrontal cortex (Abutalebi et al., 2008; Jeon & Friederici, 2015). Here, participants were learning to associate new words to novel referents, which can also be considered as a conceptual learning task. Interestingly, both associative and concept learning processes rely on the activation of a complex network, including the prefrontal cortex and the hippocampus (Murray & Ranganath, 2007; Staresina & Davachi, 2009; Mack, Preston & Love, 2020). Specifically, the medial part of the PFC contributes to the formation of long-term memory traces by favoring the binding of new information into already formed memory schemas (Tse et al., 2007; van Kesteren et al., 2012).

Here, we found electrophysiological evidence for both fast and slow consolidation-related neural plastic changes as reflected in the N400, which supports the view of a stepwise emergence of new learning-based memory traces in adults, especially when no pre-existing semantic schemas could facilitate learning and consolidation processes (Havas et al., 2018). Importantly, our results also fit well with those of a recent behavioral study investigating the role of cognitive control during the early stages of word learning (Hill & Wagovich, 2020). The authors showed the important role of cognitive flexibility in deriving the meaning of new words. The results suggested that the trade-off between cognitive control and language learning mechanisms occur differently depending upon the stages of word learning (early vs. late

Study 1

consolidation). With increasing time and learning instances, gradual consolidation of new traces may lead to more robust and precise lexical memory traces of the new words and their associations being integrated into the existing semantic networks and conceptual schemes. This may, in turn, lead to lexical and semantic retrieval of new words becoming faster and more resistant to interference with the involvement of cognitive control gradually diminishing inversely to retrieval automatization (Frischkoff et al., 2010; Chein & Schneider, 2005; 2012). A recent ERP study by Elgort and colleagues (2015) provided converging results supporting this idea. They investigated contextual word learning of a non-native language (L2). When dividing their sample into fast and slow second-language learners, the results showed a frontal N400 semantic priming effect that was also delayed around 150-200 ms in the slow learners. This result highlights the effortful nature of lexical-semantic retrieval of newly learned words, especially when learning has been weak (Elgort et al., 2015). Similar results were reported after one-shot learning exposure (Mestres-Missé et al., 2007), when learning new-words based on new phonological contrasts (Dittinger et al., 2016, 2017; Rasamimanana et al., 2020) or when no stored semantic are readily available to facilitate word learning (as it is the case in the present experiment). Further studies will be needed to understand better the exact timing of modulations of N400 amplitude as well as the topographical transition from frontal to parietal locations across time, and to clarify to which extent they directly point to the increased demand of cognitive control mechanisms aiding in the effortful stages of initial word learning.

3.5 Conclusions

We provide combined behavioral and electrophysiological evidence for the dynamics of acquiring new lexical information during encoding and consolidation, from the early stages of learning until the late phase after consolidation took place. Our data show the stepwise emergence of newly learned lexical memory traces for which previously stored semantic schemas cannot facilitate learning and consolidation processes. These results favor an integrative view of language learning that involves a continuous and weighted interplay between language-specific and domain-general cognitive control brain networks.

Study 1

Acknowledgments

We kindly thank all the participants for their engagement in the present study, as well as the funding agencies for helping us carry out. This work was supported and co-funded by the BrainTrain Research Center of Excellence at the Åbo Akademi University (2015-2018, grant to M.L.) and the Ministerio de Economía, Industria y Competitividad, through the project PSI2015-69132-P to C.F. and the project PGC2018-099859-B-I00 to A.R.F. Moreover, M.L. was supported by funds from the Academy of Finland (Grants No. 260276 and 323251).

Study 1

TABLES

| Factor | Statistic | Factor | Statistic |
|------------------------|--|--|--|
| Day (1, 24) | $F = 9.89$ $p = .004$ $r^2 = .29$ | Day x Anteropost (2, 48) | $F = 13.62$ $p = .001$ $r^2 = .36$ |
| Block (2, 48) | $F = 14.39$ $p < .001$ $r^2 = .38$ | Block x Anteropost (4, 96) | $F = 2.11$ $p = .117$ $r^2 = .08$ |
| Day x Block (2, 48) | $F = 14.49$ $p < .001$ $r^2 = .38$ | Day x Block x Lat (4, 96) | $F = 5.92$ $p = .001$ $r^2 = .20$ |
| Day x Lat (2, 48) | $F = 2.10$ $p = .145$ $r^2 = .08$ | Day x Block x Anteropost (4, 96) | $F = 4.78$ $p = .013$ $r^2 = .17$ |
| Block x Lat (4, 98) | $F = 3.39$ $p = .023$ $r^2 = .12$ | Day x Block x Lat x Anteropost (8, 192) | $F = .78$ $p = .577$ $r^2 = .03$ |

Repeated measures ANOVA *Experiment 1*

Table 2. Statistics from the repeated measures ANOVA of Experiment 1 for the N400 (350-550 msec) with 4 factors: Day (Day 1, Day 5), Block (Block 1, Block 2, Block 3), Lateralization (Left, Central, and right) and Antero-posterior (Frontal, Central, and Parietal). Only the main effects of the experimental manipulations and the main effects per electrode factors are reported here, thus, no main effects of electrode factors (Lateralization & Position) are shown. *Lat = Lateralization.

| Factor | Statistic | Factor | Statistic |
|-------------------------------|--|--|---|
| Group (1,18) | $F = .02$ $p = .881$ $r^2 = .001$ | Group x Block x Lat (4, 144) | $F = 1.23$ $p = .302$ $r^2 = .03$ |
| Block (2, 72) | $F = 17.72$ $p < .001$ $r^2 = .33$ | Block x Anteropost (2, 48) | $F = 2.09$ $p = .133$ $r^2 = .06$ |
| Group x Block (2, 72) | $F = 3.98$ $p = .024$ $r^2 = .10$ | Group x Block x Anteropost (4, 144) | $F = 1.99$ $p = .144$ $r^2 = .05$ |
| Group x Lat (2, 72) | $F = .44$ $p = .614$ $r^2 = .01$ | Group x Lat x Anteropost (4, 144) | $F = .37$ $p = .738$ $r^2 = .01$ |
| Group x Anteropost (2, 72) | $F = .39$ $p = .585$ $r^2 = .01$ | Block x Lat x Anteropost (8, 288) | $F = .443$ $p = .678$ $r^2 = .01$ |
| Block x Lat (2, 72) | $F = 7.62$ $p < .001$ $r^2 = .18$ | Group x Block x Lat x Anteropost (8, 288) | $F = 1.62$ $p = .200$ $r^2 = .04$ |

Mixed model ANOVA *Experiment 2*

Table 3. Statistics from the mixed model ANOVA of Experiment 2 for the N400 (350-550 msec) with Group (LG vs. NLG) as between-subject factor and Block (Block 1, Block 2, and Block 3), Lateralization (Left, Central, and Right) and Antero-posterior (Frontal, Central, and Parietal) as within-subject factors. Only the main effects of the experimental manipulations and main effects per electrode factors are reported here, thus, no main effects of electrode factors (Lateralization & Position) are shown. *Lat = Lateralization.

Chapter 4

Study 2

This study corresponds to:

Neus Ramos-Escobar, Emma Segura, Guillem Olivé, Antoni Rodríguez-Fornells, Clément François. Oscillatory activity and functional connectivity of concurrent word segmentation and meaning mapping in 9-year-old children

Chapter 4 Study 2

Oscillatory activity and functional connectivity of concurrent word segmentation and meaning mapping in 9-year-old children

When learning a new language, one must segment words from continuous speech and associate them with meanings. These complex processes can be boosted by attentional mechanisms triggered by multi-sensory information. Previous electrophysiological studies suggest that brain oscillations are sensitive to different hierarchical complexity levels of the input, making them a plausible neural substrate for speech parsing. Here, we investigated the functional role of brain oscillations during concurrent speech segmentation and meaning acquisition in sixty 9-year-old children. We recorded EEG while children performed an audio-visual SL task during which they were exposed to a learning speech stream with consistent word-picture associations and a random stream with inconsistent word-picture associations before being tested on their ability to recall words and word-picture associations. We capitalized on the brain dynamics to align neural activity to the same rate as an external rhythmic stimulus to explore modulations of neural synchronization and long-distance connectivity during multi-sensory word learning. Results showed enhanced power at both word- and syllabic-rate and increased connectivity between frontal and occipital regions in the learning, compared to the random condition. These findings suggest that multi-sensory cueing and attentional mechanisms play an essential role in children's successful word learning.

Keywords: Children, speech segmentation, meaning mapping, EEG, frequency-tagging, functional connectivity.

Study 2

4.1 Introduction

When learning a new language, one must segment words from continuous speech and associate them with possible meanings. Recent cognitive models suggest that these processes may rely on SL, a domain-general mechanism that allows the computation of statistical regularities within and across sensory modalities (Frost et al., 2015; Räsänen & Rasilo, 2015). In the case of speech, infants, children, and adults can implicitly segment continuous streams of pseudo-words by relying on the TP between consecutive syllables (Cunillera et al., 2009; François et al., 2013; 2017a; Saffran et al., 1996). Redundant multi-sensory cues can facilitate both speech segmentation (Cunillera, 2010a; 2010b; Glicksohn & Cohen, 2013; Thiessen, 2010) and word-referent associations (Yu & Smith, 2007; Shukla et al., 2011; Yurovski et al., 2012). Interestingly, these redundant cues may capture learners' attention by highlighting temporally contiguous target words and referents (Axelsson et al., 2012; Horst & Samuelson, 2008). These results are in line with previous studies proposing that the combination of attentional and associative processes may facilitate multimodal word learning (Plunkett, 1997; Smith, 2000). The attention toward a visual event may trigger the association of the visual referent with the spoken word. This attentional capture may occur through cue facilitation, where the learner uses the referent's perceptual saliency as a cue for the binding process. Instead, this process could rely on associative mechanisms that allow the binding of temporally contiguous multi-sensory events (Smith, 2000). In any case, enhancing visual attention to a referent either by perceptual saliency or by temporal contiguity may facilitate word learning. Overall, current empirical evidence and theoretical proposals converge on the idea that providing redundant cues might facilitate language learning.

Several studies have used EEG to decipher the neural substrates of speech segmentation in adults (Batterink & Paller, 2017; Cunillera et al., 2009; de Diego-Balaguer et al., 2007; François et al., 2014; 2017b). Despite the importance of understanding how multimodal language learning operates, only one study has explored the neurophysiological correlates of parallel speech segmentation and meaning-mapping in adults (François et al., 2017b). In this study, participants were exposed to an audio alone speech segmentation task and an audio-visual meaning-mapping task. After exposure to random and statistically structured streams, participants were tested on the online detection of mismatches, the explicit recall of words, and word-to-picture associations. Behavioural data showed better word recognition in the audio-visual compared to the audio condition. Furthermore, the analysis of event-related brain

Study 2

potentials (ERPs) revealed a larger fronto-central N400 for structured compared to random streams in both conditions. Crucially, topographical differences between conditions suggested that additional resources over frontal regions might be necessary for the meaning-mapping.

To the best of our knowledge, no previous study has focused on brain oscillations during multimodal language learning in children. This is important as brain oscillations play a crucial role in a variety of cognitive operations such as stimulus processing (Thut et al., 2011; Worden et al., 2000), cognitive control (Canolty et al., 2006; Cunillera et al., 2012), and the development of abstract reasoning or verbal working memory (Power et al., 2012; Embury et al., 2019; Taylor et al., 2020). Interestingly, synchronization among brain regions is associated with learning (Fries, 2005) and, the intrinsic hierarchical organization of brain oscillations allows optimizing stimulus processing (Lakatos et al., 2005). Moreover, one influential tenant is that speech processing occurs by synchronizing neural activity to the speech signal (Giraud & Poeppel, 2012; Luo et al., 2010; Luo & Poeppel, 2007). Better alignment with external stimuli predicts speech comprehension and better audio-motor response (Assaneo et al., 2018; Peelle et al., 2013). Further research shows that neural synchronization to the stimulus only occurs for attended stimuli (Ding & Simon, 2012; Lakatos et al., 2013), whereas when the tracking of the stimulus is not possible, speech perception and comprehension is impoverished (Ahissar et al., 2001, Luo & Poeppel, 2007). Those findings confirm that top-down attentional mechanisms modulate speech processing in children (Wang et al., 2014; Panda et al., 2020; Ríos-López et al., 2020). One way to investigate oscillatory brain activity is to use frequency-tagging analysis, which shows the fine neural synchronization in response to both auditory and visual stimuli (Doesburg et al., 2012; Kim et al., 2007; Keitel et al., 2014; Covic et al., 2017; Nozaradan et al., 2012). This method has been used to show that successful speech segmentation is associated with a peak at both syllable and word frequencies during exposure to statistically structured streams (Batterink & Paller, 2017; Buiatti et al., 2009; Farthouat et al., 2017; Henin et al., 2020; Ordin et al., 2020). However, EEG-based studies in infants and children are still scarce (François et al., 2013, 2017a; Kabdebon et al., 2015; Teinonen et al., 2009), and no previous study has used frequency-tagging combined with connectivity analyses to explore multimodal word learning in children.

Here, we explored concurrent speech segmentation and meaning mapping in a large group of sixty 9-year-old children. We recorded EEG during an audio-visual SL task. We applied frequency-tagging and connectivity analyses to track the online modulations in neural synchronization and long-distance connectivity during learning. Connectivity analyses used the weighted Phase-Lag Index (wPLI), a measure of phase synchronization between two signals that avoids the problem

Study 2

of volume conduction (Vinck et al., 2011). Specifically, children were presented with a learning stream with consistent word-picture associations and a random stream with inconsistent associations. Based on previous results in adults (François et al., 2017b), the multimodal presentation of words and referents may facilitate word segmentation and the binding to its referent. The presence of this visual referent should induce attentional capture and facilitate the emergence of new multi-sensory memory traces. Therefore, we expected larger neural responses at both syllable and word frequencies over frontal regions for learning than for random streams. We also expected an enhanced word- frequency response over occipital regions reflecting pictures' presentation at the same frequency as words. Finally, we predicted larger crosstalk between visual and speech-segmentation brain regions with stronger connectivity between fronto-occipital regions in the learning than in the random condition.

4.2 Methods

4.2.1 Participants

A total of seventy-two children attending the fourth grade in two elementary schools in Barcelona participated in the study. Two children did not finish the task and ten more were excluded due to noise in the recordings because of too many movements (4), drowsiness (3), or background noise (3). Hence, a total of sixty children (mean age: 9 years and 5 months; 34 girls) were included in the study. The children were recruited from two elementary schools in Barcelona. All children had similar socioeconomic status as determined from the parents' incomes (middle/upper class). All the children were Catalan-Spanish bilinguals and were right-handed, except for four left-handed. Written informed consent was obtained from all the children's parents before the start of the study. School directors and teachers were also informed and collaborated in the organization of the experiment. The study was conducted following the local norms and guidelines for the protection of human subjects and was approved by the local ethics committee.

4.2.2 Stimuli

We used the same audio-visual artificial language streams as those used in François and colleagues (2017b). Briefly, we created a synthetic speech stream that contained four trisyllabic pseudo-words concatenated in a pseudo-random order with no immediate repetition. We used

Study 2

the MBROLA speech synthesizer with the Spanish male database (es1) to create the stream. All phonemes had the same duration (116 ms) and pitch (200 Hz). Each word had a duration of 696 ms and was repeated 25 times in the stream. The language stream was presented four times with a small pause between each presentation for a total of 400 words and a duration of 4 min and 39 s. The transitional probability between syllables was 1.0 within words and .33 between words.

In addition to the statistically structured stream, a random stream was created by pseudo-randomly mixing the syllables. The random stream had the same duration as the structured one, but the TP between syllables was .09. Therefore, EEG signals from this condition could be used as a baseline.

Four different pictures were synchronously presented with the four words of the structured speech stream. The visual stimuli consisted of four 20×120 mm black-and-white drawings belonging to four different semantic categories (animals, vegetables, vehicles, and accessories; Snodgrass & Vanderwart, 1980). Importantly, each pseudo-word was associated with a single picture resulting in four fully consistent word-picture associations. In the random condition, four different pictures were pseudo-randomly presented on the screen every 696 ms. However, because the random stream did not contain pseudo-words, no association could be created.

4.2.3 Experimental Procedure

Children were tested individually in a quiet room of their school in 2 separate sessions on 2 different days that included neuropsychological assessments and electrophysiological tests, respectively. The electrophysiological session lasted 1h and a half, and the neuropsychological session lasted 30 minutes.

In the neuropsychological session, children were evaluated with sub-tests of the NEPSY-II battery (Korkman et al., 2007) and the WISC-IV (Weschler, 2015). Specifically, we assessed language functions using the verbal fluency and the non-word repetition task from the NEPSY-II battery. We evaluated non-verbal problem solving with the matrix-reasoning subtest and working memory with the digit span subtest from the WISC-IV. Overall they performed within the normal range (Digit span: $M = 15.56$, $SD = 2.89$; Matrix reasoning: $M = 17.9$, $SD = 3.71$; Non-word repetition: $M = 41.18$, $SD = 2.65$; Verbal fluency: $M = 33.28$; $SD = 6.33$; Letter fluency: $M = 15.02$, $SD = 6.05$).

Study 2

During the electrophysiological session, children sat on a comfortable chair and performed different computer-based tasks while we recorded EEG. For each stream, children were explained to carefully listen with the task of discovering the words of an “alien” language. First, the random condition was presented, followed by the learning condition. After the learning phase, children were assessed with both a speech segmentation and a word-to-picture association task (**Figure 29**).

The level of performance for word segmentation was assessed with a lexical decision task. In each of the 32 trials, children heard a word from the language or a non-word. The non-words were compiled by mixing the syllables of the words pseudo-randomly. Children had to decide whether the item was a word from the language or not. We chose to assess word recognition with a lexical decision task (LDT) to collect behavioral responses for both types of test items as done in adults (François et al., 2017b).

Children’s word to picture association performance was assessed with a 12-trial associative word-picture matching task. On each trial, children heard a word of the language while two pictures were displayed on the computer screen. They had to choose which picture (left or right) was associated with the word. Importantly, both pictures appeared in the continuous streams. In both tasks, each trial was presented on the screen until participants responded, forcing them to answer before moving to the next trial.

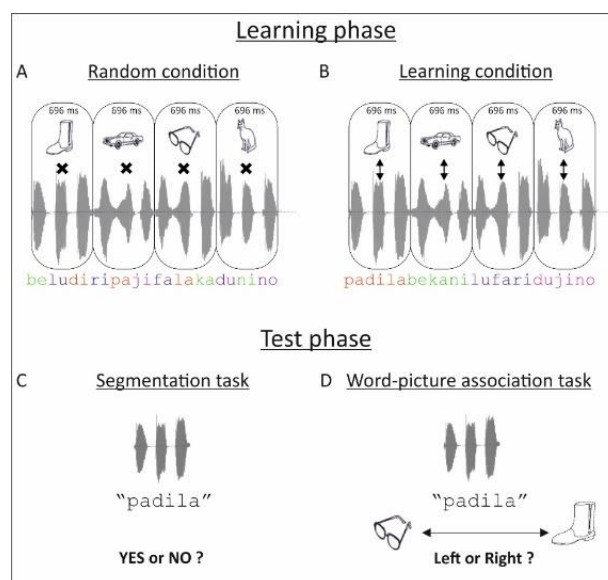


Figure 29. A random and a learning stream (A and B respectively) were used. The learning phases were immediately followed by the test phase. C: Illustration of one trial from the segmentation task. D: Illustration of one trial from the word to picture association task.

Study 2

4.2.4 EEG data acquisition and analyses

The EEG signal was recorded at a sampling rate of 256 Hz using a Biosemi amplifier system (Biosemi ActiveTwo, Amsterdam University) from 16 active Ag-Cl scalp electrodes mounted on a child-sized elastic cap at standard positions (International 10/20 system sites: Fp1, Fp2, F3, F4, T7, C3, C4, T8, P3, P4, O1, O2, Fz, Cz, Pz, and Oz). Five additional electrodes were placed on the right and left mastoids for off-line re-referencing, at 1 cm from the right and left outer canthi and below the right eye to control for blinks and lateral eye-movements.

Frequency tagging analysis of spectral power

Continuous EEG recordings were re-referenced off-line to the average of the left and right mastoids and high-pass filtered at 0.1 Hz to remove slow drifts. Epochs of 5,57 sec corresponding to the presentation of 8 words were extracted to obtain a high frequency-resolution and were baseline corrected. This large time-window provided a frequency resolution for the Fourier transform of each trial equal to 0.12 Hz. This resolution allowed reasonable estimations at both target frequencies (word frequency: 1.11 Hz and syllable frequency: 4.31 Hz). Epochs with activity exceeding baseline level by $\pm 80 \mu\text{V}$ in the [-50 700 ms] period were excluded from the analyses. We used only one of every four epochs in the average to avoid oversampling. We then transformed the obtained waveforms into the frequency domain by using a discrete Fourier transformation (Matlab; Natick, MA). Finally, the power at both target frequencies was obtained for each child and electrode in both random and learning conditions. All the scripts were manually prepared for this analysis. For the statistical analysis, we used a repeated-measures ANOVA with four within-subject factors: Frequency (Syllable Frequency, Word Frequency), Condition (Learning, Random), region of interest (ROI; Frontal, Central, Parietal, Occipital), and Lateralization (Left, Central, Right).

Connectivity analysis – weighted Phase-Lag Index

To assess the phase relationship between different brain regions during speech segmentation, we used wPLI. We selected this measure because it avoids volume conduction since zero-lag synchronization is left out from the analysis, and thus, the reported measure is not based on common sources (Bastos & Schoffelen, 2016; Stam et al., 2007). We performed the same pre-processing as for the frequency-tagging analysis, we computed the Fourier transform, and finally, the wPLI using Fieldtrip Toolbox (Oostenveld et al., 2011) running in Matlab. To calculate the wPLI, we used electrodes as nodes and computed differences (t -values) in connectivity

Study 2

between learning and random conditions, applying False-Discovery Rate (FDR) corrections when necessary. For the statistical analysis, we defined four different regions of interest (ROI): frontal (F3, F4, Fz), central (C3, C4, Cz), parietal (P3, P4, Pz), and occipital (O1, O2, Oz) to perform a repeated-measures ANOVA with three within-subjects factors: Frequency (Syllable Frequency, Word Frequency), Condition (Learning, Random), and ROI (Fronto-Occipital, Fronto-Central, Fronto-Parietal, Occipito-Central, Occipito-Parietal, Parietal-Central). We used the BrainNet Viewer to visualize significant networks (Xia et al., 2013).

4.3 Results

4.3.1 Behavioral results

Figure 2 shows the behavioral data obtained in the speech segmentation and word-picture association tasks. For the speech segmentation task (**Figure 30A**), the comparison of performance against chance level (50%) showed that children performed significantly above chance level for both words ($M = 63.4$, $SEM = 2.15$) and non-words ($M = 62.2$, $SEM = 1.78$, both P 's $< .001$). This result indicates that they were able to recall newly segmented words correctly and to reject non-words correctly. The level of performance for words and non-words was not significantly different ($t(59) = 0.35$; $P = .72$). For the word-picture association task (**Figure 30B**), the mean percentage of correct responses ($M = 70.49$, $SEM = 2.77$) was significantly above chance level ($t(59) = 7.39$; $p. < .001$) indicating that children were also able to associate the pictures to the segmented words.

Study 2

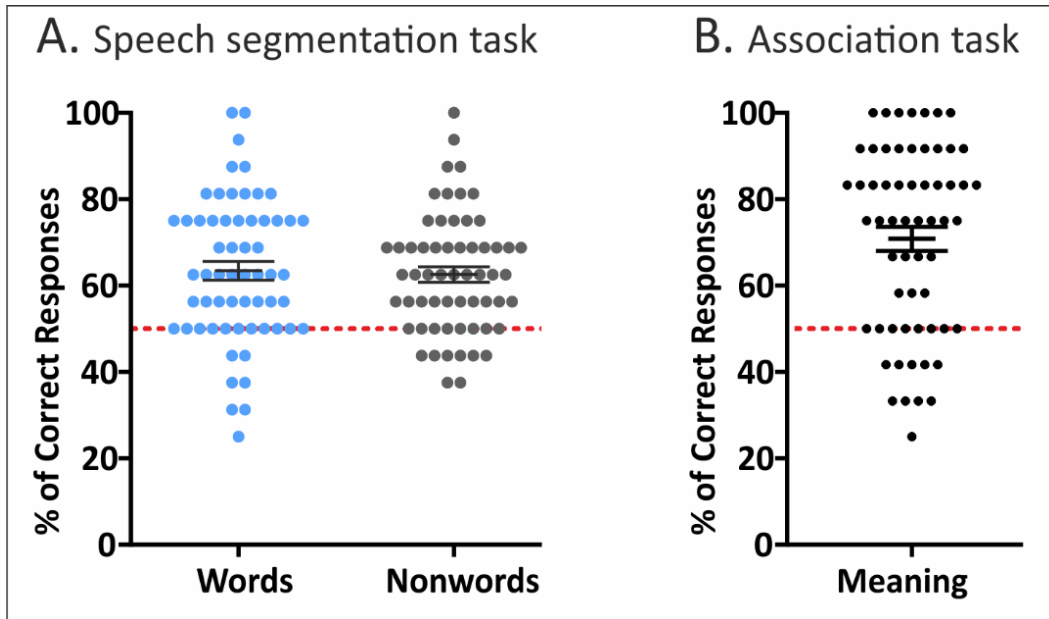


Figure 30. Behavioral data. Percentage of correct responses in the speech segmentation task for words and non-words (A) and in the word-picture association task (B). Dots represent individual values, and bars correspond to the mean and standard error of the mean (SEM) in each condition. The red dotted line represents the chance level (50%).

4.3.2 Electrophysiological results

Frequency tagging analysis

Figure 31A and **Figure 31B** display the results of the frequency-tagging analysis showing clear peaks at word and syllable frequencies. The peak to word frequency was higher than for syllable frequency [significant main effect of Frequency: $F(1,59) = 11.49$, $p = .001$, $r^2 = .16$]. As predicted considering the rhythmic presentation of pictures (matching word length), we observed a clear peak at the word frequency over occipital ROI [significant Frequency x ROI interaction [$F(1,177) = 22.09$, $p < .001$, $r^2 = .27$]. Interestingly, some ROIs showed larger differences between Learning and Random conditions than others at both word and syllable frequency [evidenced by a non-significant main effect of Condition [$F(1,59) = 2.94$, $p = .092$, $r^2 = .05$] but a significant Condition x ROI interaction [$F(3,177) = 9.18$, $p < .001$, $r^2 = .14$] and a significant Condition x Frequency x ROI x Lateralization interaction: [$F(6,354) = 4.65$, $p = .007$, $r^2 = .07$]; **Figure 31A & Figure 31B**]. To further explore these effects, we performed separate repeated-measures ANOVAs for both word and syllable frequency.

At the word frequency, a clear peak was found over all ROIs with a strong response over occipital regions [main effect of ROI: $F(3,177) = 18.14$; $p < .001$, $r^2 = .22$]. When looking at differences

Study 2

between conditions across ROIs [significant Condition x ROI interaction: $F(3,177) = 3.54$; $p = .029$, $r^2 = .06$] we found that those differences occurred over frontal regions only [$t_{59} = -2.31$; $p = 0.025$] with an enhanced power in the learning compared to the random condition. Interestingly, despite the clear peak at occipital ROI, there was no differences between conditions in this region [$t_{59} = 1.17$; $p = 0.247$]. The Condition by ROI by Lateralization interaction was also significant [$F(6,354) = 3.74$; $p = .016$, $r^2 = .15$]. Further post-hoc analysis at electrode level showed significantly enhanced power at word frequency in the learning compared to the random conditions at F4, Fz, F3 and P3 electrodes (F4: $t_{59} = -2.2$; $p = .032$; Fz: $t_{59} = -2.01$; $p = .048$; F3: $t_{59} = -2.45$; $p = .017$; P3: $t_{59} = -2.18$; $p = .033$; see **Figure 31C**). However, none of these post-hoc comparisons survived the FDR correction for multiple comparison.

At the syllable frequency, we found clear peaks in all ROIs with a stronger response over frontal regions [main effect of ROI: $F(3,177) = 6.51$; $p = .002$, $r^2 = .10$]. Results also revealed an overall enhanced power in the learning compared to the random condition [main effect of Condition: $F(1,59) = 6.81$; $p = .01$; $r^2 = .10$]. Interestingly, this effect was differently distributed across ROIs [significant Condition x ROI interaction: $F(3,117) = 15.99$; $p < .001$, $r^2 = .21$] with enhanced power in the learning compared to the random condition over frontal ($t_{59} = -4.59$; $p < .001$) and central regions ($t_{59} = -3.44$; $p = .01$). No differences between learning and random were found over parietal ($t_{59} = -.64$; $p = .526$) and occipital ($t_{59} = 1.17$; $p = .247$) regions. The Condition by ROI by Lateralization interaction was also significant [$F(6,354) = 3.41$; $p = .011$, $r^2 = .06$]. Further post-hoc analysis at electrode level showed that the power at the syllable frequency was larger in the learning condition than in random (FDR correction, $p < .004$) at F4, Fz, F3, C3, Cz, C4 electrodes (F4: $t_{59} = -4.66$, $p < .001$; Fz: $t_{59} = -4.11$, $p < .001$; F3: $t_{59} = -4.68$, $p < .001$; C3: $t_{59} = -3.77$, $p < .001$; Cz: $t_{59} = -3.11$, $p = .003$; C4: $t_{59} = -3.02$, $p = .004$, see **Figure 31D**).

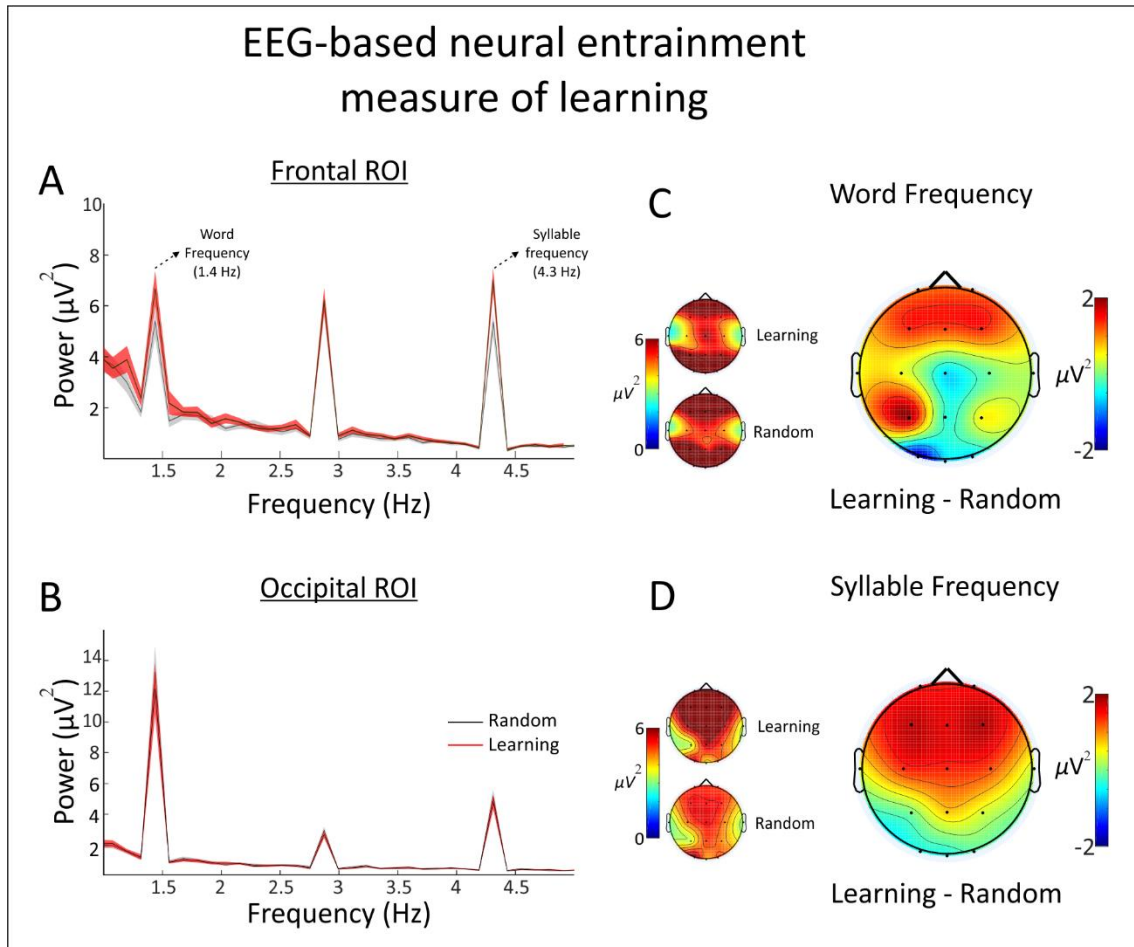


Figure 31. Results of the frequency tagging analysis. Power spectrum of the EEG signal over frontal (A) and occipital (B) regions of interest (Mean and SEM). Red = learning condition, black = random condition. Topographical distribution of the power at syllable (C) and word frequency (D) for the single conditions and the difference between conditions (Learning-Random).

Connectivity analysis

Based on the frequency-tagging results, we expected a larger cross-talk between visual and speech-segmentation brain regions during the learning condition, reflecting the involvement of visual cues in word binding and integration. Specifically, we expected stronger fronto-occipital connectivity at word frequency. **Figure 32A** & **Figure 32B** show the significant connections between electrodes at word and syllable frequency, respectively. The wPLI values were stronger at the syllable than at the word frequency [significant main effect of Frequency: $F(1,59) = 72.91$; $p < .001$, $r^2 = .55$]. On **Figure 32C** the connectivity spectra across ROIs are depicted, pointing to a higher connectivity strength in the learning compared to the random condition [main effect of Condition: $F(1,59) = 15.69$; $p < .001$, $r^2 = .21$]. As can be seen on **Figure 32D** & **Figure 32E**, this effect was differently distributed between brain regions [significant Frequency x Condition x ROI

Study 2

interaction: $F(1,59) = 3.26$; $p = .021$, $r^2 = .05$]. Further post-hoc t -tests conducted at the word frequency

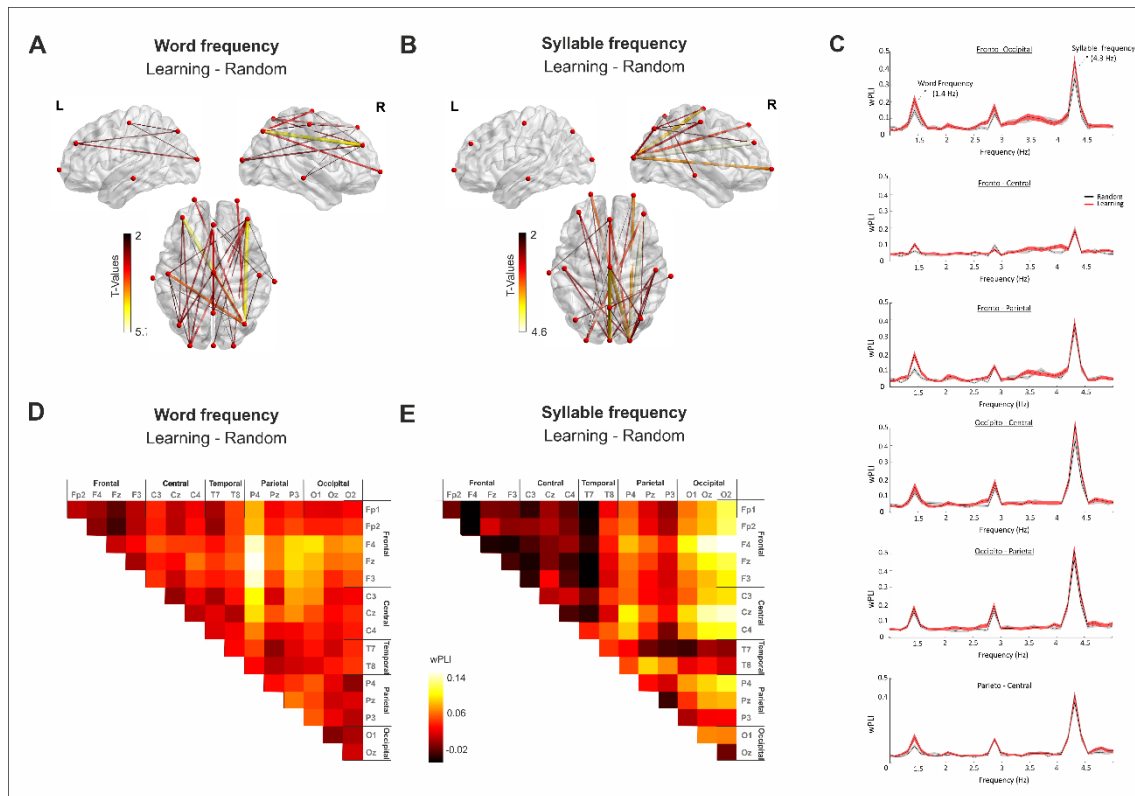


Figure 32. Results of the connectivity analysis. Statistical connectivity maps for the learning - random contrast at the word (A) and the syllable frequency (B). The color scale indicates t-values with light colors indicating strong connectivity differences. C) Connectivity spectra between the different ROIs (mean and SEM). Red = learning condition, black = random condition. D & E) Connectivity matrices at the word (D) and syllable (E) frequency for each pair of electrodes. The color scale indicates wPLI values, with light colors indicating strong connectivity.

showed stronger connectivity in the learning compared to the random condition in fronto-occipital ($t_{59} = 2.91$, $p = .005$), fronto-central ($t_{59} = 2.77$; $p = .008$), fronto-parietal ($t_{59} = 4.38$, $p < .001$) and parieto-central networks ($t_{59} = 2.9$; $p = .005$; FDR correction, $p < .008$; see **Figure 32D**). For the syllable frequency we found significantly stronger connections in the learning compared to random condition in fronto-occipital ($t_{59} = 3.38$; $p = .001$) and occipito-central networks ($t_{59} = 2.98$; $p = .004$; FDR correction, $p < .004$; see **Figure 32E**).

4.4 Discussion

In the present study, we investigated the functional role of brain oscillations during concurrent speech segmentation and meaning acquisition in sixty 9-year-old children. Specifically, we used frequency tagging and connectivity analyses to assess neural tracking and long-distance

Study 2

connectivity during an audio-visual SL task. Behavioural results showed that children successfully learned the novel word-forms and their associated meaning. Frequency-tagging results revealed enhanced power at word and syllable frequencies in the learning compared to the random condition. Furthermore, we found evidence of enhanced long-distance connectivity between frontal and parieto-occipital regions at both target frequencies. These converging results suggest that attentional mechanisms modulate word segmentation and meaning mapping during children's multimodal word learning.

At the behavioural level, the results showed that children's performance was above chance level for both the speech segmentation and the word-to-picture association tasks. These results are the first showing that 9-year-old children can successfully perform speech segmentation and meaning mapping concurrently. This finding is in line with previous reports showing that redundant cues such as prosody can trigger word-object associations in infants (Plunkett, 1997; Smith, 2000; Shukla et al., 2011). The present results also replicate our previous study conducted in adults in which the presence of visual referents resulted in refined memory traces of the correct syllabic patterns (François et al., 2017b). Our results suggest that children and adults exhibit similar behavioural patterns of explicit recall.

The frequency tagging analysis revealed clear peaks at both syllable and word frequencies during exposure to the streams. At both target frequencies, we found enhanced power in the learning compared to the random condition. However, the differences survived FDR correction only for the syllable frequency suggesting that differences between conditions were higher when tracking the syllables than the words. Importantly, the level of neural alignment to the speech streams was significantly stronger at the word than at the syllable frequency. This result confirms previous findings in adults showing a temporal alignment of brain activity with speech and musical stimuli (Luo et al., 2010; Luo & Poeppel, 2007; Nozaradan et al., 2012). Here, we focused on EEG raw power and observed enhanced activity in the learning condition compared to the random at the word- and syllable frequency. While some studies have also focused on raw power as we did here (Buiatti et al., 2009; Nozaradan et al., 2014), others have used inter-trial phase coherence (ITC) where the phase is also considered (Buiatti et al., 2009; Batterink & Paller, 2017; Henin et al., 2020). Results based on ITC analyses showed higher coherence in the learning than in random at word frequency but an opposite pattern at syllable frequency with higher ITC values in the random than in the learning condition. Several reasons may account for those differences between previous studies and the present.

Study 2

First, methodological differences may explain the different patterns of results observed. A recent study used both measures to explore the oscillatory activity during an auditory SL task and found a similar pattern of results (Ordin et al., 2020). While ITC provides information about the precision of the synchronization across trials, raw power gives information about the strength of the brain response. Therefore, although ITC and power measures are related, they provide different, not overlapping information (Wöstmann, Fiedler & Obleser, 2017). Second, our study is the first exploring oscillatory activity in children during speech segmentation. Thus, maturational brain differences between children and adults may also explain those differences. For instance, children's neuronal alignment to speech differs from adults. A recent study found that cortical tracking of speech at the syllable rate (4-8 Hz) predicted speech-in-noise discrimination ability in 6- to 9-year-old children. Notably, the detrimental effect of increasing auditory noise on the level of neural tracking was stronger in children than in adults suggesting that children may exhibit an immature cortical tracking of speech (Vander Ghinst et al., 2019). Third, Hollich, Hirsh-Pasek & Golinkoff (2000) proposed that the relative weight of linguistic and non-linguistic cues in word learning facilitation changes over development. Specifically, while children may be biased to rely on domain-general cues during the initial stages of word-learning, they may use language-specific information (e.g. grammatical rules) during later learning stages. Thus, differences with adults may account for developmental changes in speech-brain tracking and attentional factors such as the reliance and weighting of different cues (Rios-Lopez et al., 2020; Moore et al., 1999). Here, visual cues may have facilitated the discovery of words by prompting participants to focus on the first syllable of the words (synchronized with the picture presentation), which may, in turn, modulate sensory processing (Parasuraman, 1980; Woldorff et al., 1993). The high response to word-frequency found in occipital regions further supports this idea (see **Figure 32B**). Because the pictures were presented at the same frequency as the words, this neuronal response over occipital regions presumably reflects the neural alignment to pictures. Steady-state visual evoked potentials studies consistently showed that attention toward the stimulus could modulate the neuronal response to that stimulus (Hillyard et al., 1998; Keitel et al., 2013; Morgan et al., 1996; Kim et al., 2007). Therefore, the attentional capture by visual referent may enhance sensory processing and facilitate the whole word learning process.

Connectivity analyses based on the wPLI revealed enhanced long-distance connectivity strength in frontal-occipital and frontal-parietal networks at word frequency in the learning compared to the random condition. This result suggests that attentional modulation induced by consistent

Study 2

visual cues may contribute to facilitating the learning. Previous studies using similar methodologies have shown that the neural tracking of a specific talker in higher-order auditory regions differs between attended non-attended speech input (Ding & Simon, 2012; Lakatos et al., 2013). Moreover, endogenous fluctuations of brain excitability may drive differences in behavior during an attentional task (Helfrich et al., 2018). These reports suggest that voluntary attention to a specific stimulus modulate the neural tracking and, therefore, stimulus encoding. Studies on infant's gaze-following showed that word learning differences arise from different object-looking behaviors (Yu & Smith, 2012). Besides, multiple forms of visual attention may differently support statistical word-referent learning (see Talsma et al., 2010 for a review; Smith, 2000). Therefore, the fronto-occipital connectivity patterns found here may reflect attentional modulations arising from visual cues and targeting frontal and temporal auditory speech brain regions. Those attentional modulations may be crucial for optimizing multimodal word-learning mechanisms (Smith, 2000; Plunket, 1997; Hollich, Hirsh-Pasek & Golinkoff, 2000). However, the spatial resolution of the 16 electrodes EEG system used here is somewhat limited. Further studies with a better spatial resolution must be conducted to understand better the brain mechanisms involved in multi-sensory word learning.

Finally, an important concern about neural oscillations is whether external stimuli directly induced them by phase-resetting or whether they merely reflect highly phase-locked brain response to the repeated stimulus (see Obleser & Kayser 2019 for a discussion). Some researchers tried to discern these issues by using indirect measures such as sub-threshold stimulation (Oever et al., 2017) and found that oscillation activity still arises in this condition. Others studied the oscillation activity while after the stimulus presentation based on the idea that an oscillator would continue for a while after the presentation of the stimulus (Lakatos et al., 2013). Recently, Doelling and colleagues (2019) showed that the phase was more variable in an evoked model than in an oscillation model. This finding, thus, supports the hypothesis that brain oscillations arise from synchronization with external stimuli and endorse frequency tagging as a tool to measure neuronal tracking of external stimuli in children.

4.5 Conclusion

To our knowledge, this is the first study investigating the functional role of brain oscillations and long-distance connectivity during multimodal SL in children. Our results show (i) that 9-year-old children can simultaneously segment words and associate them onto meaning, (ii) that the presence of multimodal cues triggers speech segmentation and meaning mapping process in

Study 2

children, (iii) that long-distance connectivity between visual and speech areas contribute to multimodal word-learning. These results favor an integrative view of language learning, where multi-sensory cues and attentional mechanisms interact to enhance learning.

Acknowledgments

We thank all the children and their families for participating in the study. We also thank the directors and the teachers of the two schools where the children were tested. We thank D. Cucurell, G. Cardonna, and H. Tajadura for their help in different steps of the study. This research has been supported by a Spanish MINECO project (PSI2015-69132P) to CF.

Chapter 5

Study 3

This study corresponds to:

Neus Ramos-Escobar, Manuel Mercier, Agnès Trébuchon-Fonséca, Antoni Rodríguez-Fornells, Clément François, Daniele Schön. Hippocampal and auditory contributions to language SL.

Chapter 5 Study 3

Hippocampal and auditory contributions to speech segmentation

Statistical learning has been proposed as a mechanism to structure and segment the continuous flow of information in several sensory modalities. Previous studies proposed that the medial temporal lobe and in particular the hippocampus may be crucial to parse the stream in the visual modality. However, the involvement of the hippocampus in auditory statistical learning and specifically in speech segmentation is less clear. To explore the role of the hippocampus in speech segmentation based on statistical learning, we exposed seven pharmaco-resistant temporal lobe epilepsy patients to a stream composed of trisyllabic pseudowords and recorded intracranial stereotaxic electroencephalography (sEEG). We used frequency-tagging to assess the neuronal synchronization of the hippocampus and auditory regions to the temporal structure of words and syllables of the stream. Results show that while auditory regions highly respond to syllable frequency, the hippocampus responds mostly to word frequency. This finding provides direct evidence of the involvement of the hippocampus in the speech segmentation process and further suggests a hierarchical organization of auditory information during speech processing.

Keywords: Hippocampus, SL, frequency-tagging, sEEG, speech

Study 3

5.1 Introduction

Humans are daily exposed to a massive amount of information. Finding a structure in these massive sensory inputs is an efficient way to make sense of the world. Structural information can be decoded thanks to regularities in the input that can be tracked by computing low-order statistics (Reber, 1967, Frost et al., 2015). This ability is often referred to as statistical learning (SL), a domain-general learning mechanism that has been described across motor (Hunt & Aslin, 2001), visual (Fiser & Aslin, 2002a) or auditory (Saffran, et al., 1996, 1999) modalities.

SL has been proposed to be a possible mechanism that allows segmenting word units from fluent speech (Cutler & Butterfield, 1992; Saffran et al., 1996). This process seems to occur incidentally and without effort, via simple exposure (Saffran et al., 1997; Turk-Browne et al., 2005; Saffran et al., 1999). Word segmentation is relevant for the isolation and identification of word units in continuous speech. Indeed, speech segmentation is one of the first problems that language learners must deal with when learning a new language (Graf-Estes et al., 2007; François et al., 2017b). Although several behavioral (Cutler & Butterfield, 1992; Saffran et al., 1996) and electrophysiological studies (Sanders et al., 2002; Cunillera et al., 2006; de Diego-Balaguer et al., 2007; Ablamkin et al., 2008; Schön et al., 2008; François et al., 2014; 2017) have been conducted to characterize the process, the underlying precise brain network dynamics are not yet clear.

In the first study in SL carried out with fMRI, McNealy and colleagues (2006) found that middle and superior temporal regions (STG) and IFG were activated during SL (see also Ablamkin et al., 2008; Turk-Browne et al., 2009, Cunillera et al., 2009; Karuza et al., 2013; Schapiro et al., 2014). Because SL occurs across different modalities (Henin et al., 2020), previous research has addressed the differential implication of sensory areas in the computation of regularities according to the stimulus modality. Studies of SL with visual stimuli showed activations in visual networks including the lateral occipital cortex, the occipito-temporal cortex, and the parieto-occipital sulcus (Turk-Browne et al., 2009; Bischoff-Grethe et al., 2000) whereas experiments using auditory stimuli resulted in activation of primary and secondary auditory regions (i.e. STG or planum temporale; McNealy, Mazziotta, & Dapretto, 2006; Cunillera et al., 2009; Karuza et al., 2013). Importantly, previous fMRI studies showed that not only cortical structures are involved in SL but also the hippocampus (Turk-Browne et al., 2009). The interplay between cortical and subcortical structures in SL fits well with cognitive models that propose learning and memory process to occur in two different stages, from an initial encoding of the information supported by MTL and hippocampal structures to long-term storage of episodic memory traces mediated

Study 3

by neocortical regions (Davis & Gaskell, 2009; McClelland et al., 1995). Notwithstanding the consensus on the involvement of classical language regions, few studies have focused on the role of hippocampal regions in language processing (Piai et al. 2016; Henin et al., 2020).

The hippocampus plays a crucial role in segmenting continuous sensory inputs into discrete events (Zacks & Swallow, 2007; Radvansky & Zacks, 2017). Recent studies on event memory formation propose that the interplay between sensory regions and the hippocampus supports the creation of boundaries between events. Specifically, while sensory areas seem to be responsible for fine-grained boundaries, the hippocampus instead supports cortical information binding into memory traces (Baldassano et al., 2017; Ben-Yakov & Dudai, 2011; Zacks et al., 2001; Speer et al., 2007). In addition, recent studies on vocabulary acquisition based on associative or contextual learning consistently show functional activations of the hippocampus during the early stages of learning (Bartolotti et al., 2017; Breitenstein et al., 2005; Covington & Duff, 2016; Mestres-Missé et al., 2008; Ripollés et al., 2016; Züst et al., 2019). Nonetheless, direct human electrophysiological evidence for the role of the hippocampus in extracting pattern regularities found in speech is still missing.

Oscillations between frequency ranges are associated with multiple cognitive processes and, by interacting with each other, regulate the flow of information in the brain (Buzsáki & Dragun, 2004; Canolty et al., 2006). In the hippocampus of rats, theta-gamma cross-frequency coupling (CFC) is proposed to mediate item-context associative learning and memory recollection processes (Tort et al., 2009). Further research in humans showed that theta oscillations can also modulate the amount of contextual linguistic information to facilitate language processing (Piai et al., 2016). Here, we used intracortical recordings to assess the involvement of the hippocampus in speech segmentation based on SL. We used frequency-tagging analysis to obtain a measure of neural entrainment of both auditory and hippocampal regions to the stimuli (Buiatti et al., 2009; Ding et al., 2016; Batterink et al., 2017). In a recent study, Henin and colleagues (2020) gathered intracortical brain responses from human epileptic patients during an auditory SL task. Using ECoG they mapped the cortical circuits for SL and found STG to be involved in earlier processing stages (e.g. syllables), whereas IFG was involved in higher-order units (e.g. words). However, they were not able to assess the involvement of the hippocampus in the extraction of linguistic regularities.

Here, we used intracranial recordings from 7 patients with pharmaco-resistant temporal lobe epilepsy implanted with depth electrodes in the auditory cortex and the hippocampus during an

Study 3

auditory SL task. We used frequency-tagging to assess neural entrainment of auditory and hippocampal regions to the frequency of the syllables and words. We expected auditory regions to show a peak in the power spectrum corresponding to the frequency of the syllable (phonological processing), while the hippocampus would respond to the frequency of words (word segmentation).

5.2 Methods

5.2.1 Participants

Seven patients with pharmaco-resistant temporal lobe epilepsy (4 females, mean age = 29; range 18-45) participated in the study (see **Table 4**). Patients were implanted with depth electrodes for clinical reasons to determine the epileptic zone before they underwent neurosurgical treatment at the Hospital de La Timone in Marseille (France). The location of the implanted electrodes was solely determined by clinical criteria. Patients provided informed consent prior to the experimental session, and the study was approved by the Institutional Review Board of the French Institute of Health (IRB00003888).

| Patients | Gender | Age (years) | Hemispheric dominance | Epileptogenic zone | Depth electrodes | Hippocampal electrodes |
|----------|--------|-------------|-----------------------|----------------------|------------------|------------------------|
| P1 | F | 29 | L | L temporal | 4R + 10L | Both |
| P2 | F | 45 | L | R temporal | 10R + 2L | Both |
| P3 | F | 18 | L | R temporal | 5R + 4L | Both |
| P4 | F | 23 | Atypical | L temporal | 1R + 12L | L |
| P5 | M | 19 | L | L temporal | 2R + 11L | R |
| P6 | M | 42 | L | L Frontal | 1R + 13L | L |
| P7 | M | 33 | L | R Frontal & Parietal | 14R | R |

Table 4. Patients clinical description. *M* male, *F* female, *L* left, *R* right

5.2.2 Data acquisition & electrode localization

The sEEG signal was recorded using depth electrodes of 0,8 mm of diameter containing 10 or 15 contacts (Alcis, Besançon, France). The contacts were 2 mm long and were spaced from each other by 1.5 mm. Data was recorded using a BrainAmp amplifier system (Brain Products GmbH,

Study 3

Munich, Germany), sampled at 1000 Hz, and high-passed filtered at 0.016 Hz. During the acquisition, recordings were referenced to a single scalp-electrode located at Cz. Channels were offline converted to a bipolar montage (closest neighbor electrode reference) to increase the spatial resolution, SNR, and to remove common sources due to neighbor areas (Mercier et al., 2017).

To precisely localize electrode contacts, a procedure similar to the one used in the iELVis toolbox was applied (Groppe et al., 2017). First, we manually identified the location of each contact centroid on the post-implant CT scan using the Gardel software (Medina et al., 2018). Second, we performed cortical reconstruction and volumetric segmentation on the pre-implant MRI with the Freesurfer image analysis suite (documented and freely available for download online <http://surfer.nmr.mgh.harvard.edu/>). Third, we mapped electrode locations to the pre-implant MRI brain (processed with FreeSurfer) and to the MNI template, using SPM12 methods (Penny et al., 2006). Statistical parametric mapping. *Anal. Funct. Brain Images*, (2006), through the FieldTrip toolbox (Oostenveld et al., 2011). The co-registration to the patient brain was done via a rigid affine transformation to respect individual anatomy. The normalization to the MNI template was done through a non-linear transformation to map electrodes to a standardized space and allow brain regions labeling using the Destrieux atlas (Destrieux et al., 2010).

5.2.3 Experimental procedure

We used a similar experimental design to the one used in our previous studies with healthy adults and children (Schön et al., 2008; François & Schön 2011; François et al., 2013; 2014). Briefly, it consisted of an implicit learning phase followed by an explicit recognition 2-alternative forced-choice test. During the implicit learning phase, patients were asked to listen carefully to an auditory stream without explicit instructions of learning. Immediately after the learning phase, patients performed the behavioural test. In each trial of the test phase, they had to press one of two buttons to indicate which of 2 words (first or second item) most closely resembled what they had just heard in the continuous stream (see **Figure 33**). In each test trial, one word was a “legal” auditory word-form present in the stream while the other item was a “nonword” created by pseudo-randomly mixing the syllables of the words. Thus, these nonwords were “illegal” items that were never heard before. Each familiar word of the language (legal word) was presented with each unfamiliar word (nonwords), making up 16 pairs that were repeated twice, leading to 32 test trials.

Study 3

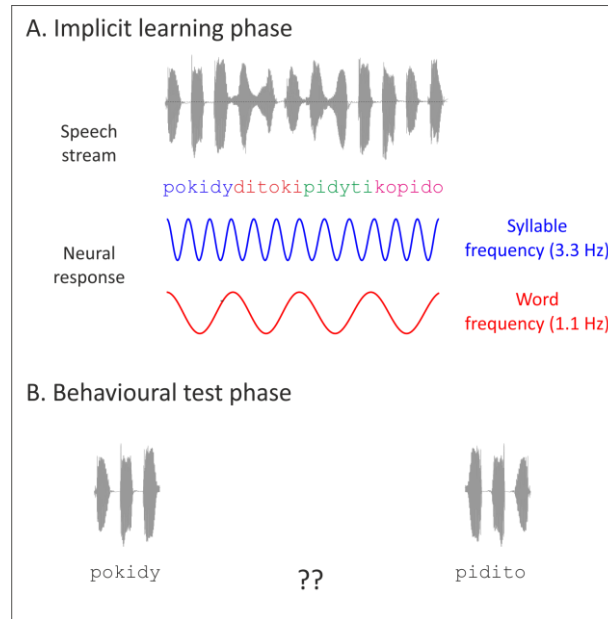


Figure 33. Illustration of the experimental procedure. After exposure to a continuous stream of syllables, participants performed a 2-alternative forced-choice task to assess the level of learning.

5.2.4 Stimuli

The language consisted of 4 consonants ('p', 't', 'k', 'd') and 3 vowels ('o', 'i', 'y'), which were combined into a set of 11 syllables. The exact syllable length was set to 300 ms. These syllables were then combined to give rise to 4 tri-syllabic spoken words (POKIDY, DITOKI, PIDYTI, and KOPIDO) and then synthesized using Mbrola (<http://tcts.fpms.ac.be/synthesis/mbrola.html>). More precisely, the speech stream was built by concatenating 7 minimal sequences of non-coarticulated syllables respecting the constraint of not repeating the same word twice in a row. Importantly, no acoustic cues have been inserted at word boundaries. Each word was presented 30 times thus leading to a stream of 120 words, lasting 1 min and 48 seconds. For the test phase, the 4 nonwords used were TOPIDY, DYPOKI, KOKITI, and PIDITO.

5.2.5 sEEG Data analysis: Frequency tagging

For each patient, sEEG data were visually inspected using AnyWave software (Colombet et al., 2015) and channels with artifacts or with epileptic seizures were excluded from the analysis. Continuous sEEG recordings acquired during the learning task were filtered using a 0.5 Hz high pass filter to remove slow drifts in the recorded signal. Then, epochs time-locked at the onset of each word were created by segmenting the recordings from 4 words before and 4 after the stimulus yielding epochs of 8 words length (lasting 7.2s). Epochs were partially overlapping, but

Study 3

we took care to use an overlap equal to twice the size of the word in order to ensure that possible artifacts would not lead to a spurious peak at the word frequency. A baseline correction of 4 words was applied. Epochs were averaged and transformed to the frequency domain using a discrete Fourier transformation (Matlab; Natick, MA). We extracted the power values for each target frequency (word frequency: 1.11 Hz; syllable frequency: 3.33 Hz; two-syllables frequency: 1.67 Hz). Power values at the target frequencies were obtained for each patient and electrode.

5.2.6 Statistical analysis

We first assessed whether hippocampal contacts showed significantly greater responses to word-frequency than all the electrodes located elsewhere in the brain. For each patient and each target frequency (word, syllable & two syllables), we computed the distribution of power values across contacts. As expected, the distribution was not normal. We, therefore, selected a non-parametric threshold (median, Mdn, plus 2.5 interquartile range, IQR) to determine the significant contacts (see **Figure 35**). The criterion for defining hippocampal contacts was done by visual inspection of the anatomical data (see **Figure 34**) and confirmed via automatic atlas labeling (see above). The same procedure was also applied to electrodes in the primary auditory cortex (henceforth auditory electrodes). Whenever more than one contact was present in the same region, the average was used. We then normalized the data and computed the difference between hippocampal and auditory electrodes for each patient at both word frequency, two-syllable frequency, and syllable frequency and applied the Wilcoxon test to assess whether the difference was different from zero.

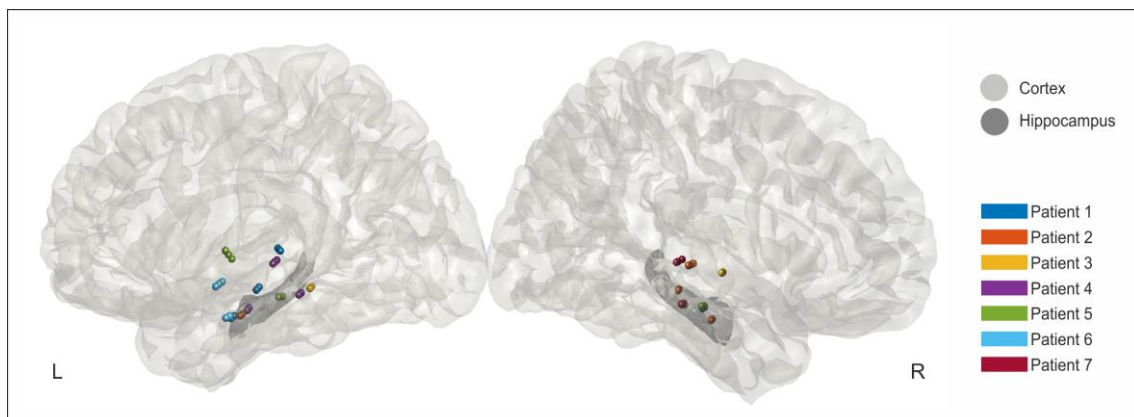


Figure 34. sEEG electrode location. Colored dots indicate the electrode location for each patient in the auditory and hippocampal regions. Clear grey represents the cerebral cortex and dark grey depicts the hippocampus.

Study 3

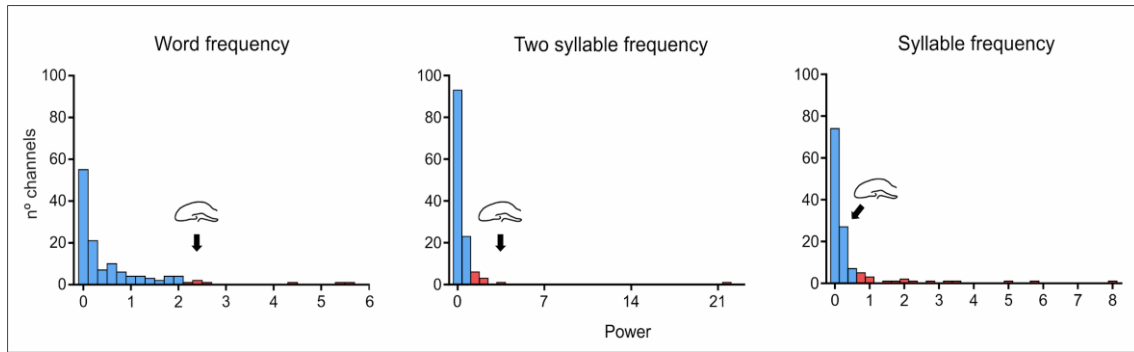


Figure 35. Example of the methodology used to define hippocampal implication. Histograms of power response to word, two-syllable, and syllable target frequencies, for Patient 6. Red bars indicate the values above the median plus 2.5 IQR. Black arrows indicate the power values at hippocampal electrodes (in this example significant at the word and two-syllable frequencies).

5.3 Results

5.3.1 Behavioural results

The level of performance in the 2AFC test showed that the percentage of correct explicit word recognition was 37,5% (range: 25-56%) confirming previous results of impaired explicit word recall in patients with epilepsy (Schapiro et al., 2014; Henin et al., 2020).

5.3.2 Electrophysiological results

Clear power spectrum peaks at word and syllable frequencies are visible over auditory and hippocampal contacts (**Figure 36A**).

For the syllable frequency, all patients except one exhibited a clear peak in contacts located within the primary auditory cortex (Mdn = 12.24; IQR = 315.69). Five out of six patients also showed significant responses at this target frequency in hippocampal contacts although smaller than auditory responses (Mdn = 1.62, IQR = 2.76).

For the word-frequency, all patients except one (Patient 4) showed a significant response in hippocampal contacts (Mdn = 3.86; IQR = 15.95). Three patients (Pat 1, 3 & 5) also showed a significant response to word-frequency in auditory contacts although smaller than hippocampal responses (Mdn = 1.62; IQR = 8.73).

Study 3

For the two-syllable frequency, all patients showed a significant response at hippocampal contacts (Mdn = 4.79 IQR = 5.87). By contrast, none of the patients showed a significant response to the two-syllable frequency in auditory contacts (Mdn = 0.59; IQR = 0.71).

The amplitude of the peaks in the power spectrum significantly differed in hippocampus and auditory regions across all target frequencies (word frequency: Mdn = 0.98, $p = .01$; two-syllable frequency: Mdn = 1.97, $p = .01$; syllable frequency: Mdn = -2.15, $p = .03$).

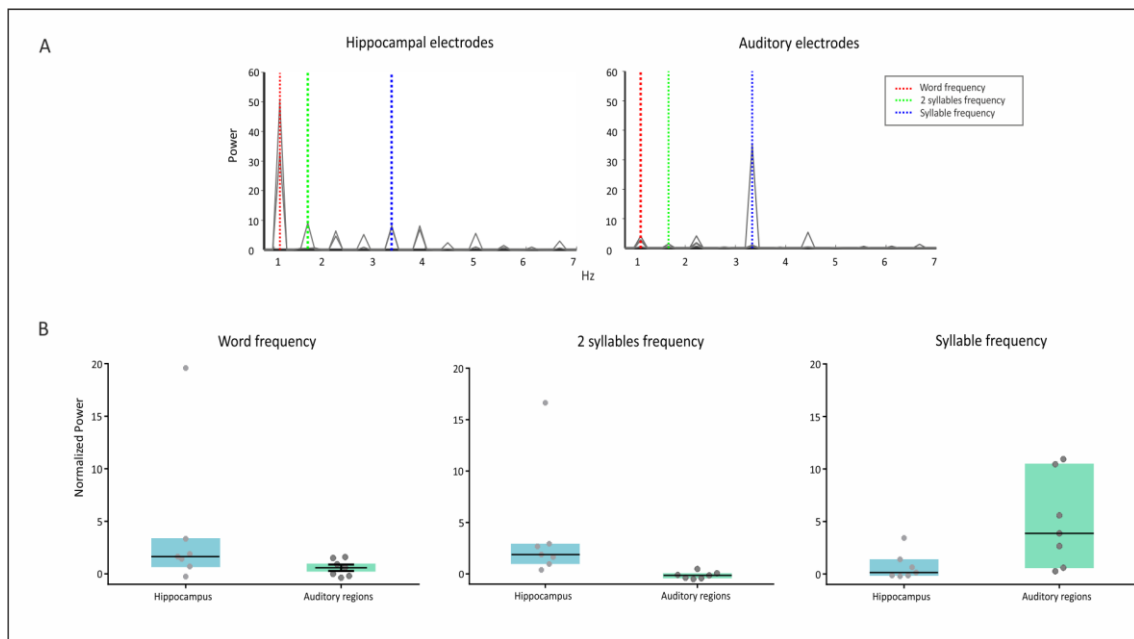


Figure 36. A) Example of a patient (Patient 7) power response of hippocampal and auditory electrodes to word frequency (red), two-syllable frequency (green), and syllable frequency (blue). B) Average of all patients' neural responses to word, two-syllables, and syllable frequencies in the hippocampus and auditory regions (Z-score-normalized data). Black lines indicate the median of all patients and box plots indicate the IQR.

5.4 Discussion

In the present study, we directly assessed the role of auditory regions and the hippocampus during speech segmentation based on SL. Pharmaco-resistant epileptic patients implanted with sEEG depth electrodes in both regions listened to a continuous stream of statistically organized syllables. Frequency tagging analysis showed that the hippocampus preferentially responded to the word-frequency rate while auditory regions preferentially responded to syllable frequency rate instead (see **Figure 36B**). Although previous studies have suggested the possible involvement of MTL regions and especially the hippocampus in SL, here we provide direct

Study 3

electrophysiological evidence for the role of the hippocampus during speech segmentation based on SL.

It is interesting to appreciate that our results clearly differ from previous fMRI studies on speech segmentation, in which no activation has been reported in MTL structures during SL (McNealy Mazziotta, & Dapretto, 2006; Cunillera et al., 2009; Karuza et al., 2013). However, the present results are also consistent with neuropsychological studies suggesting that the hippocampus and MTL structures are associated with the extraction of regularities (Schapiro et al., 2014) and previous neuroimaging SL studies in the visual modality (e.g., a sequence of shapes or scenes; Turk-Browne et al., 2009; Schapiro et al., 2012). The involvement of the hippocampus during SL in the visual modality is not surprising due to the importance of this region in visuospatial memory (Cohen & Eichenbaum, 1993; Ryan et al., 2000; Davachi, 2006). The most compelling evidence supporting the importance of the hippocampus in SL comes from a single case study in which a patient with a complete loss of hippocampus and extensive damage surrounding MTL performed an SL task in the visual (using shapes and scenes) and auditory (tones and syllables) modalities (Schapiro et al., 2014). This patient was unable to learn pattern regularities in any of these two modalities. Notice that an intend of replication of this study s showed that some forms of SL could occur without its direct participation. Covington and colleagues (2018) replicated the case study with a larger sample size and found that patients performed above chance level in the SL tasks, although they were overall impaired in comparison to healthy controls. In a follow-up study, three of those patients performed a cross-situational learning task (Warren et al., 2019). Word-referent mapping through cross-situational learning is supposed to rely on SL mechanisms (Yu and Smith, 2007; see for recent evidence, Peñaloza et al., 2017). The findings in Warren and colleagues' study pointed out to the same idea as the previous study (Covington and colleagues, 2018): two of the three patients performed above chance level in the cross-situational learning task but below the performance observed in healthy controls. These findings suggest that despite the hippocampus participates in the neural network sustaining SL, its participation might not be strictly necessary.

Overall, the present results are in convergence with recent proposals highlighting the possible contribution of the hippocampus and MTL regions in language processing (see Duff & Brown-Schmidt, 2012). For example, in a recent study, Piai and colleagues (2016) found an increase of theta oscillations in the hippocampus during the comprehension of high constrain context sentences, that is, sentences containing final words that could be easily deduced by the sentence

Study 3

context. These results suggest a crucial role of the hippocampus during language comprehension and the reactivation of lexical information from memory.

Importantly, we found that the hippocampus responded to the word frequency but also to the two-syllable frequency. This finding is in agreement with the idea that the hippocampus is sensitive to pattern regularities presented in the environment. A possible explanation of this finding is that learners may use a clustering strategy to extract the regularities (Swingley, 2005). That is, learners would tend to relate and group sounds or syllables that co-occur. This has been demonstrated in experiments showing that infants cluster the syllables that have high co-occurrence probabilities (Aslin, Saffran and Newport, 1998; Saffran, Aslin & Newport, 1996). In those experiments, as in the present, stimuli consisted of trisyllabic words. Therefore, within words, the second syllable is always preceded by the first and, at the same time, is always followed by the third. In these situations, it seems reasonable to think that the hippocampus is also sensitive to the co-occurrence of syllable duplets within the words (even the smaller statistical units, duplets in the present case). Thus, one possible function of the hippocampus during SL could be to participate in storing discretized events, creating boundaries between them, and forming associations of consecutive pairs of syllables that eventually would form the possible triplets to be stored. This idea fits well with the computational model of clustering (PARSER, Perruchet & Vinter, 1998) that proposes that continuous speech is segmented by creating smaller chunks and that those temporal chunks or templates are maintained when they often co-occur. Some models further proposed (Baldassano et al., 2017) that the small chunks are encoded in primary sensory regions whereas long events encoding would occur in higher-level brain regions including cortical regions (Zacks et al., 2001; Speer et al., 2007) and the hippocampus (Ben-Yakov & Dudai, 2011; see also Zacks & Swallow, 2007 & Radvansky & Zacks, 2017).

Several caveats could limit the interpretation of our results and might be important to acknowledge. First, the electrophysiological modulations observed in the hippocampus and auditory cortex were not reflected at the behavioral performance, and patients performed poorly in the explicit recognition test. A possible explanation of this dissociation may be that the behavioral test may not be sensitive enough to assess the level of learning (François, Tillmann & Schön, 2012; Batterink et al., 2015). Alternatively, the electrophysiological results may be reflecting implicit learning that cannot be consciously retrieved or it is still not possible to access consciously this recently learned information. Furthermore, we cannot exclude the possibility that brain damage or seizure location may explain the lack of behavioral evidence. Convergent

Study 3

to our results, Henin and colleagues (2020) did not find evidence for a correspondence between electrophysiological and behavioral responses either. These authors argued that online measures such as electrophysiological measures might be more sensitive to reflect implicit learning and SL than offline measures. In the same direction, a recent study showed that the hippocampus is sensitive to pattern regularities despite participants were not consciously aware of those regularities (Turk-Browne et al., 2010). This could explain the lack of correspondence between behavioral and electrophysiological results.

Second, out of seven patients, only one (Patient 4) did not show a significant response to word frequency in the hippocampus contacts. Further clinical exploration showed that this patient had an atypical functioning of the hippocampus, possibly due to the disease-related history. Previous studies with patients showed that hippocampal damage could lead to impairments not only in declarative memory (Squire et al., 2004) but also in SL tasks (Kim et al., 2009; Turk-Browne et al., 2005; Schapiro et al., 2014; Covington et al., 2018). Likely, the low hippocampal response of this patient might be driven by an alteration of the normal hippocampal function.

Overall, our results provide preliminary direct evidence for the involvement of the hippocampus in the brain network that orchestrates auditory speech segmentation based on SL. The results further suggest a hierarchical organization of the auditory information during speech processing. While primary auditory areas would be responsible for the extraction of low-order auditory information (the syllables), the hippocampus would be involved in the representation of pattern regularities among the lower-level of information (the possible target words). Replication of the present results in a larger sample is mandatory in order to confirm our findings.

5.5 Conclusions

Here, we directly assessed the role of the hippocampus in speech segmentation based on SL. We showed a high level of synchronization of the hippocampus to word frequency but not to syllable frequency. Conversely, auditory regions highly responded to syllable frequency but not to word frequency. This finding provides preliminary evidence in humans for the involvement of the hippocampus in speech segmentation based on SL brain networks. These results suggest a hierarchical organization of auditory information during speech processing where both cortical and hippocampal regions contribute to language learning.

Study 3

Acknowledgments

We kindly thank all the patients and their families for participating in the study. We also want to thank Patrick Marquis for his kind help and collaboration in the project.

Chapter 6

General Discussion

Chapter 6 General Discussion

Learning a language is a complex process that encompasses many high-level cognitive abilities, including perception, attention, and cognitive control (Fitch, 2010). A fundamental aspect of language learning is vocabulary acquisition, a process by which learners integrate novel words into the mental lexicon. How adults and infants acquire vocabulary has been a significant matter of interest for decades. Several neural mechanisms have been proposed to explain how word learning occurs in both L1 and L2 learning (Rodríguez-Fornells et al., 2009; Bloom, 2000; Tomasello, 2003). ***The present dissertation aimed to expand the knowledge in the field by studying the mechanisms involved in language learning in both adults and children.*** Specifically, the main objectives of this dissertation were i) to explore the temporal dynamics of word learning, ii) better understand how multimodal information is used to extract and associate meaning to novel words in parallel, and iii) directly evaluate the role of the hippocampus in the speech segmentation process.

To achieve these goals, we developed three studies exploring different aspects of word learning with EEG, sEEG, and behavioural measures. Although a detailed discussion of each study can be found at the end of each experimental section, this section provides an overview of the main findings. In the following, I discuss the implications and consider how the present results contribute to the current knowledge. Some limitations of the present studies are also considered as well as some future directions.

6.1 Summary of results

Study I examined the neural dynamics of the gradual encoding and consolidation of novel word-forms into the mental lexicon in two ERP experiments. In *Experiment 1*, 25 healthy adults learned novel objects and their meaning during a 5-day training program. In *Experiment 2*, we randomly assigned 38 new participants to a consistent learning group (Cons) or non-consistent learning group (Incons). While in the Cons, the associations word-object were consistent through the learning, in the Incons did not, thus, impeding the learning. Results of *Experiment 1* showed that successful learning was accompanied by a progressive decrease of the N400 component with a shift of the topographical distribution of the N400 from frontal to parietal after consolidation. In *Experiment 2*, we replicated this finding showing a similar decrease in the N400 amplitude for

General Discussion

the Cons group but not in the Incons group. We also found a significant correlation between the N400 and the participant's executive function abilities, thus, suggesting the involvement of cognitive control mechanisms during the early stages of vocabulary acquisition.

In **Study II**, we explored the neural synchronization mechanisms underlying speech segmentation and meaning mapping in children using a multimodal speech segmentation task. A recent study suggested that, in adults, the extraction of words from fluent speech can be facilitated by the simultaneous presentation of visual referents (François et al., 2017b). Here, we aimed to explore whether this process is similar in children and to examine the underlying neural correlates. With this aim, we exposed 60 nine-year-old participants to an audio-visual speech segmentation task while we recorded EEG. We used frequency tagging and connectivity analyses to quantify the brain response at word and syllable frequency. The results revealed that word extraction and meaning acquisition could co-occur in children during this paradigm as revealed by an above chance performance in the behavioral tests. The neural activity exhibited increased brain response to words and syllables and an increased connectivity strength between fronto-occipital regions during the learning. The results may suggest that long-distance connectivity between fronto-occipital and fronto-parietal regions mediates the learning process.

The main goal of **Study III** was to evaluate the engagement of the hippocampus in speech segmentation based on SL. Previous studies have suggested the involvement of the hippocampus in the extraction of statistical regularities in the visual modality (Turk-Browne et al., 2009; Henin et al., 2020). Nonetheless, the hippocampus's role in auditory SL and speech segmentation is less clear (Schapiro et al., 2014). Therefore, we exposed seven pharmaco-resistant epileptic patients implanted with intracranial electrodes to an auditory SL stream of trisyllabic pseudo-words. We used the frequency tagging approach to determine whether the hippocampus and auditory regions responded to syllable and word frequencies. We found that the hippocampus highly responded to words but not syllables, whereas auditory regions showed the opposite pattern. Thus, we provided direct evidence in humans for the hippocampus's involvement in speech segmentation based on SL. These findings also suggest a hierarchical brain organization for speech processing. In the following pages, I summarize and discuss the results of this dissertation in a more integrated and general manner.

General Discussion

6.2 Domain-general vs. domain-specific mechanisms

The uniqueness of human language is undoubted. However, whether language requires general or specific cognitive skills is still under debate. Over several decades, linguists and psychologists have explored the *specific characteristics* of the language system (Chomsky, 1986; Markman, 1990). More recent research has challenged this view by suggesting that language builds on a set of *general abilities* such as memory, attention, and executive functions (Bloom, 2000; Piaget, 1954; Tomasello, 1995). The debate arose because, for example, words are specifically part of the language system. Still, the sound production system or the memory mechanisms needed to retain the word are not language-specific. Indeed, the acquisition of new words may require the interplay between language-specific (Baddeley & Hitch, 1974; Gupta, 2003) and domain-general resources such as memory, attention, cognitive control, and motivation (Rodríguez-Fornells et al., 2009; Hollich, Hirsh-Pasek & Golinkoff, 2000). Which aspects of language learning are specific and which belong to general cognitive abilities are still an open question.

In the present work, we provided indirect (Study I) and direct (study II) evidence for the involvement of non-language-specific resources during the initial stages of word learning. Specifically, in **Study I**, we showed that word learning occurs in two different time scales: a fast encoding of object-words associations and a slow consolidation of new episodic memory traces in the mental lexicon. We also showed an association between the N400 modulation along with sessions and the ability in phonemic fluency. This last finding suggests a link between the neural modulations taking place during learning and verbal fluency performance. In **Study II**, we showed that multisensory cues might increase attention to enhance the segmentation and meaning association during the initial exposure to the novel word. This was evidenced by a clear increased response to pictures in the learning compared to the random condition in occipital regions and increased connectivity between fronto-occipital regions at the word frequency. These findings suggest that multimodal cues facilitate learning by capturing the learner's attention and highlighting the temporal contiguity of the stimuli.

A large body of literature attempted to solve the domain-specific versus domain-general question by exploring the brain regions involved in language learning. The history of neuropsychological research has placed language function in left-lateralized fronto-temporal regions (Broca, 1861; Catani et al., 2005; Rodríguez-Fornells et al., 2009). However, more recent neuroimaging work has extended this language network to other domain-general regions. These

General Discussion

domain-general regions include the frontoparietal control network (Vincent et al., 2008), the default mode network (Raichle et al., 2001), and subcortical areas such as basal ganglia and the hippocampus (McNealy Mazziotta, & Dapretto, 2006; Turk-Browne et al., 2009; Karuza et al., 2013). These networks co-activate with language-specific areas when performing linguistic-related tasks such as syntactic processing or word learning (Zaccarella & Friederici, 2015). However, it is still unclear whether these non-specific language regions are fully involved in the process or are merely doing an assistive function (Campbell & Tyler, 2018). In **Study III**, we provided evidence of the involvement of the hippocampus in speech segmentation based on SL. The fact that the hippocampus, a non-language-specific brain structure, is involved in the speech segmentation process suggests, again, that domain-general resources are required at least during the initial steps of language learning.

Importantly, not all aspects of language learning require the same amount of domain-general resources. While syntax seems to fit the criteria for domain-specificity (Chomsky, 1957), other aspects of language such as semantics or speech segmentation may require more general resources. For instance, complex learning situations such as noisy environments may require a higher involvement of domain-general resources such as attention and executive functions (Alain & Bernstein, 2008; Kaya & Elhilali 2017). Hauser and colleagues (2002) proposed a model that differentiated between aspects of language that are specific to language ("Narrow Language Faculty") and language in general, that is, including other cognitive abilities that are not specific to language ("Broad Language Faculty"). In their proposal, the authors claimed that language in the narrow sense is the only part of the language-specific to human beings and that the other cognitive mechanisms included in the Broad sense are shared with other animal species. This hypothesis is derived from Chomsky's theory proposing that language is a complex ability by which the human brain is highly specialized (Chomsky, 1957). In **Study I**, we found a switch of the topographical distribution of the N400 from an initial frontal distribution to a parietal region in later stages of learning. This effect may suggest that prefrontal resources related to the cognitive control process may be more involved during the initial stages of word learning. The progressive shift to parietal regions may reflect an automatization or lexicalization of the novel word forms.

Despite the findings described above, the results should be carefully interpreted as we cannot directly relate brain activations found here with specific cognitive functions (Nathan & Del Pinal, 2017; see section 6.6 for a further discussion). Overall, it appears challenging and ambitious to determine which aspects of language learning are language-specific and which are domain-

General Discussion

general mechanisms instead. Presumably, language development is a combination of both. Nonetheless, it seems relatively clear that during the initial stages of word learning, more general resources are needed (such as attention or cognitive control). In contrast, in more advanced learning stages, the process becomes more automatized, and then fewer general mechanisms are required (De Diego-Balaguer et al., 2016; Ansaldo & Enfield, 2016; Rodríguez-Fornells et al., 2009; Laine & Salmelin, 2010; Sliwinska et al., 2017).

6.3 Brain regions involved in language learning

There is a growing consensus that large-scale distributed regions support language ability. While cortical areas only were initially thought to be involved in language processing (Hickok & Poeppel, 2007), there is growing evidence suggesting that other regions may contribute to language learning (Rodríguez-Fornells et al., 2009). **Study III** provides direct evidence that auditory regions are involved in speech segmentation and that non-language specific regions such as the hippocampus participate in the process. Furthermore, in **Study II**, we found a high brain response over occipital regions and increased connectivity between fronto-occipital regions at the frequency of the word. This suggests the communication flow of information between those regions may be a basis for learning.

While classical cortical language areas would be in charge of the initial processing of the auditory signal, other regions such as memory-related or cognitive control regions may subserve and modulate the learning process. In this direction, different cognitive models have proposed that cortical regions work together with subcortical memory-related areas to facilitate word learning (Ullman, 2001; Davis & Gaskell, 2009; Rodríguez-Fornells et al., 2009). These results are also in line with research suggesting that several broadly distributed domain-general processes exert cognitive control during the learning process (see Fedorenko, 2014; for a review). fMRI word learning studies have revealed that frontal, temporal, and parietal regions co-activated during word learning (Badre et al., 2005; Badre & Wagner, 2007), adding evidence for a broader activation compared to classical language cortical network only (Lau et al., 2008).

Figure 37 summarizes some of the language-related and non-related brain regions involved in language learning explored in this work. In **Study I**, we found a change of the topographical distribution of the N400 from a frontal distribution the first day of training to more parietal regions the fifth day. This finding suggests a change of the underlying networks with the involvement of frontal sources during the initial encoding of novel word forms and the shift to

General Discussion

parietal sources reflecting the consolidation of the word forms (Koechlin, Ody & Kouneiher, 2003). In **Study II**, we showed larger responses in the occipital regions to word frequency and increased connectivity of this region with frontal areas. This finding suggests that occipital regions involved in visual processing interact with frontal language-related regions to facilitate the learning. In **Study III**, we showed that the hippocampus is directly involved in the extraction of pattern regularities during speech segmentation based on SL.

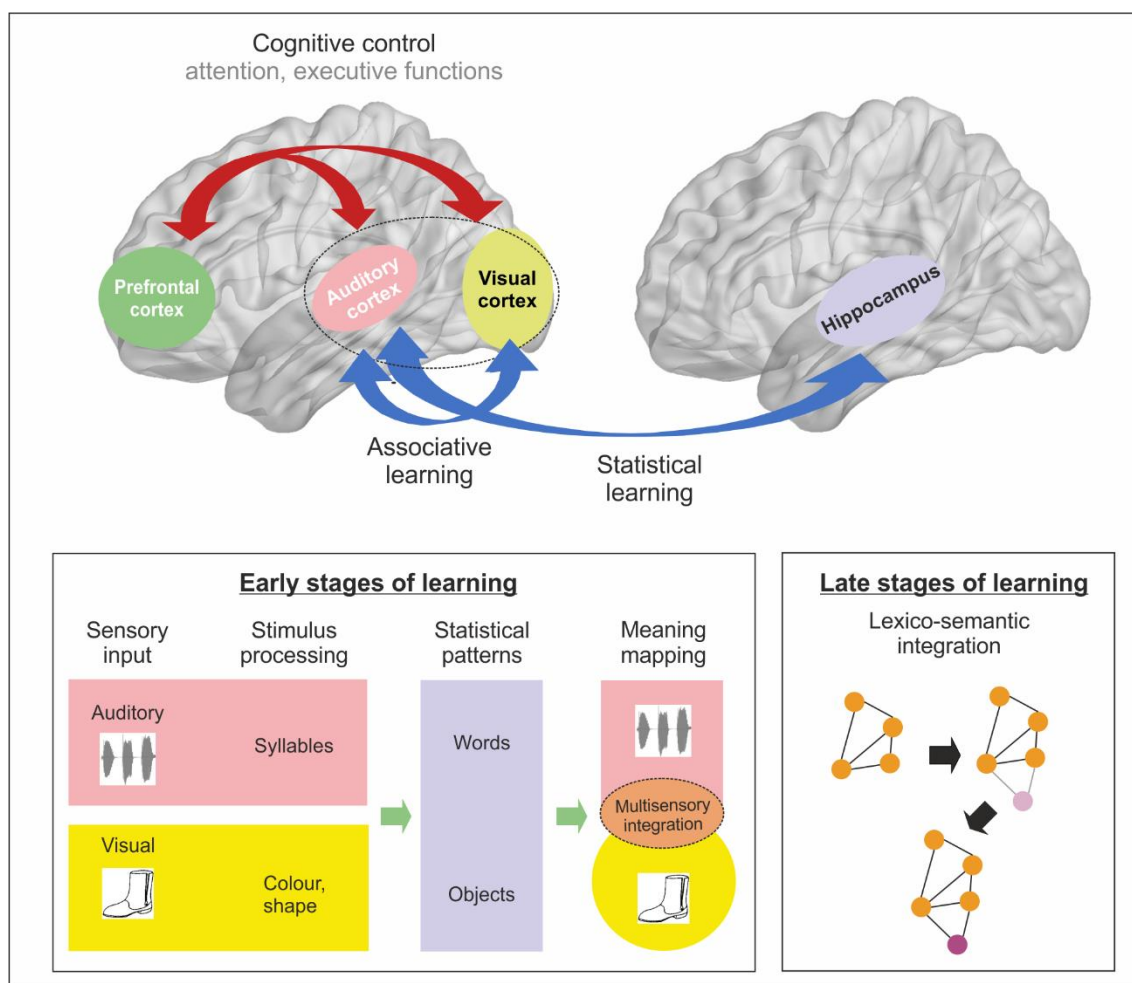


Figure 37. An integrative view of language-related brain structures and associated mechanisms involved in language learning explored in this work. In the early stages of learning, sensory inputs are initially processed by primary sensory regions such as the auditory and visual cortex. The hippocampus then would be responsible for binding patterns of information among the sensory input. Thus, this hierarchical processing would also occur during speech segmentation in which the auditory cortex may be responsible for syllabic processing while the hippocampus would be responsible for the word extraction. At the same time, this process may be facilitated by cognitive control mechanisms such as attention and executive functions. Then, the multisensory integration facilitates the association of the word form with a meaning (meaning mapping). Late stages of learning occur once meaning has been attached to the word. Then, the meaning of the novel word form (represented as the purple dot in the figure) is gradually integrated onto the existing lexico-semantic network (dark yellow network). Red arrows indicate learning mechanisms (statistical learning and associative learning); blue arrows indicate domain-general mechanisms (attention, cognitive control).

6.4 Simultaneous speech segmentation and word-referent mapping

Vocabulary acquisition entails extracting the novel word-form from fluent speech and associating it with the corresponding semantic representation (Gupta & Tisdale, 2009). In **Study II**, we showed that these two fundamental aspects of language learning could occur simultaneously in children. This finding was supported by the above-chance performance in both behavioral tasks.

Studies on speech segmentation have demonstrated that both infants and adults use statistical regularities to compute the TP between adjacent syllables to isolate the word units (Saffran et al., 1996; 1999; Cunillera et al., 2009). Nonetheless, other linguistic and non-linguistic cues such as prosody, phonotactics, or rhythmic cues (Cutler & Norris, 1988; Thiessen & Saffran, 2003) are also used for speech segmentation. When syllables are properly grouped by high-probability relationships, these lexical units can be extracted and are available to be associated with meaning (Ngon et al., 2013; Graf-Estes et al., 2007).

Likewise, to a correct mapping of words onto their referent, associative mechanisms ease to solve the ambiguity of the simultaneous facing of multiple words with multiple referents (Smith & Jones, 2011; Yu & Smith, 2007). One of the prominent theories in this area proposes that cross-situational learning (Gleitman, 1990) allows the association of novel word-forms onto its referent by accumulating statistical information of the co-occurrence of spoken words with their possible visual referents (like objects or events) across multiple encounters. Thus, optimal learning would require the gradual capture of higher co-occurrence of word and its referent (Smith & Yu, 2008). While each situation in which the word is presented may be ambiguous, by multiple encounters with similar situations and through principles of mutual exclusivity (Yu & Smith, 2012), the referent's uncertainty is reduced and gradually resolved.

Even though both speech segmentation and meaning mapping may require SL mechanisms, studies of word learning in adults have traditionally considered these two steps separately: first, the extraction of the word-form from fluent speech and later its association onto a meaning. Recent studies have shown that despite speech segmentation can occur without any other information rather than the statistical probability of syllable co-occurrence (Saffran et al., 1996; Cunillera et al., 2009) when images of the word meaning are presented together with the novel words, the learning is facilitated (Cunillera et al. 2010, François et al., 2017). In this line, Räsänen and Rasilo (2015) proposed a combined model of word segmentation and meaning acquisition

General Discussion

in which both processes are mediated by SL mechanisms. The model proposes that learners would use statistical information to segment the speech like TP between and within words but also statistical probabilities of co-occurrence between two stimuli to associate the meaning to the word-form. Hence, the simultaneous occurrence of both speech segmentation and meaning acquisition may facilitate the learning process by SL mechanisms.

An important factor for the multisensory cueing mechanism is that the co-occurrence of the stimuli must be consistent across the multiple encounters (Arnal et al., 2011). This issue is especially important in infant learning. Some literature showed that when infants consider a speaker as unreliable, the information provided by that speaker is not used to acquire the knowledge (Koenig et al., 2004; Pasquini et al., 2007). Children spontaneously keep track of a speaker's story and are more likely to learn novel words contained in the story when the speaker is accurate and consistent. However, attention to the speaker is reduced when children think they are inaccurate and, thus, are less likely to learn new words and concepts (Birch et al., 2008). In ambiguous associative contexts, the information that is considered irrelevant by the learner is filtered out while attention is driven towards the informative stimulus (Smith et al., 2014). This finding evidence that top-down attentional mechanisms highly modulate the learning process. In **Study II**, we showed an increased neural tracking of both words and syllables when the visually presented objects were consistently associated with the words (Learning condition) compared to when inconsistently associated (Random condition). Moreover, the connectivity results showed an increase of connectivity between fronto-occipital regions in the learning compared to the random condition. Importantly, in both conditions, visual cues are present but only are informative in the Learning condition. Thus, the increase of connectivity between fronto-occipital regions seems to be reflecting multisensory binding modulated by attention. Overall, the evidence suggests that multisensory integration plays a crucial role in word learning and meaning mapping process.

Although not reported in the main experiments of this dissertation, in *Experiment 1* of **Study I**, we also analyzed the ERPs in response to the pictures' presentation. The goal was to examine how the consolidation and familiarization of word-picture associations influence visual processing and further word anticipation. As can be seen in **Figure 38**, we found a clear emerging N300 to pictures that preceded the to-be-learned words during the course of the learning. The N300 is proposed to be sensitive to semantic information during object categorization tasks (Barrett & Rugg, 1990; Holcomb & McPherson, 1994). In our experiment, the N300 appeared over frontal regions during the last day of training (Day 5), while it was not visible during the

General Discussion

early training stages (Day1). We interpret this finding as an anticipatory recall performance of the word-form when the picture-word association has already been consolidated (Batterink & Neville, 2011).

We also found a modulation of the Late Positive Component (LPC) to the pictures (**Figure 38**). The LPC amplitude increased across blocks over centro-parietal electrodes during Day 1 and reached a plateau on Day 5. The LPC is a positive deflection peaking around 600 ms after stimulus presentation that has been related to the retrieval and recollection of episodic information (Rugg & Curran, 2007). The LPC has also been proposed to reflect the integration of information into working memory (Petten et al., 1991). Therefore, the increase of LPC amplitude to novel objects might be an index of robust episodic memory trace formation (Rugg et al., 1995). Taken together, our results suggest that novel objects associated with novel words might engage deeper semantic processing during later stages of training (Ganis et al., 1996; Hamm et al., 2002).

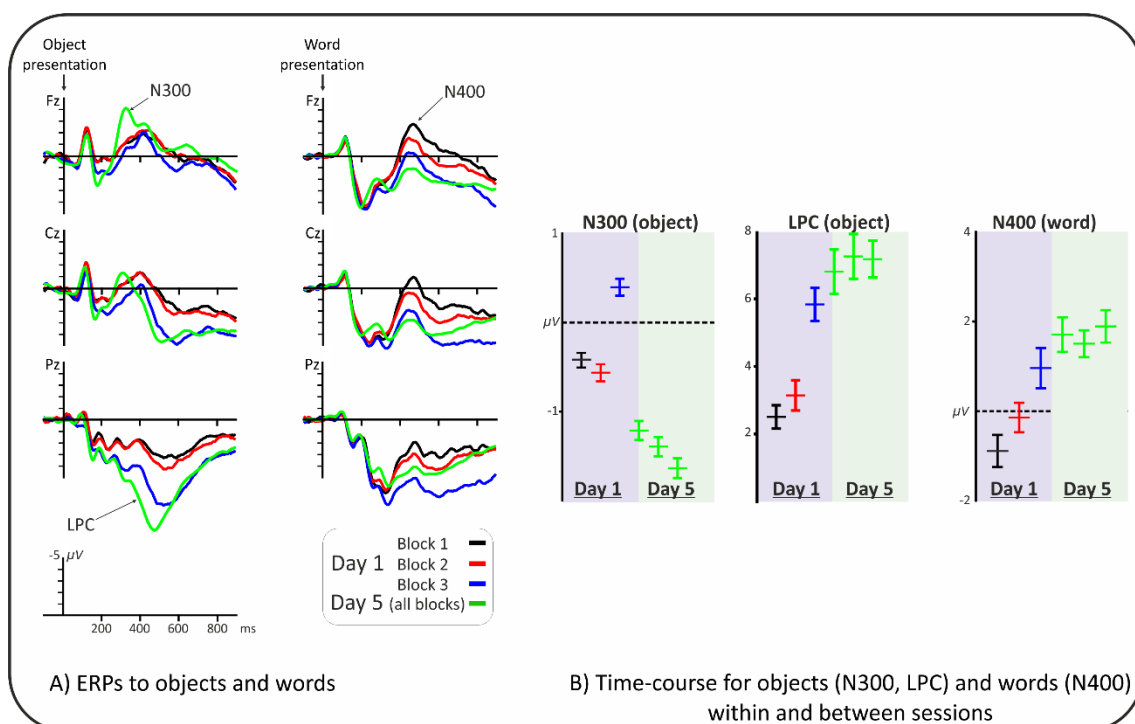


Figure 38. A) Grand averages of ERPs to object and word presentation at midline electrodes across 25 participants recorded during each learning block (black = block 1 of Day 1, red = block 2 of Day 1, blue = block 3 of Day 1, green = average of the three blocks of Day 5). B) Time-course of within- and between-session for the N300 to object presentation averaged over frontal electrodes (F3, Fz, & F4) in the 250-350 ms time-window, for the LPC to object averaged over parietal electrodes (P3, Pz, & P4) in the 450-600 ms time-window, and for the FN400 in frontal regions (F3, Fz, & F4) averaged across all blocks of Day 1 and Day 5 in the 350-550 ms time-window.

General Discussion

Importantly, to further determine the role of the N300 in lexical retrieval during picture exposure, I designed another ERP experiment. In this new experiment, participants were exposed to novel word forms and novel pictures in a five-day training program. During the learning phase, they were exposed to three types of word-pictures associations: the novel consistent (learning condition), the novel inconsistent (random), and the familiar consistent (control). We expected a modulation of the N300 amplitude when the picture-word association had been integrated into the lexico-semantic network only. Thus, by the last day of training, we expected novel objects in the learning condition to elicit an increased N300 (replicating the experiments above) compared to the objects presented in the random condition. We also expected to see N300 modulation for existing picture-word pairs since the first day of training. Nonetheless, the study was disrupted by the pandemic situation, and the experiment has not been completed yet.

Overall, the findings reported here suggest that audio-visual integration of the stimuli plays an important role in language learning, specifically in integrating and retrieving lexico-semantic information into the current lexicon. In particular, once words and pictures are already familiar, the presentation of the picture might facilitate the fast retrieval of the word-form by activating the related semantic networks.

6.5 Neuronal synchronization as a mechanism for language learning

Brain networks can synchronously oscillate at several frequency bands from 0.05 Hz to 500 Hz (Buzsáki & Dragun, 2004). However, both single neurons and networks respond preferentially to certain types of frequencies (Hutcheon & Yarom, 2000). This preference is determined by the physical properties of the neuronal membrane, such as membrane conductance and potential (Buzsáki & Dragun, 2004). The detection and response of a neuron or a network to a particular stimulus will depend on its sensitivity to the frequency and intensity of the presented stimulus (Hutcheon & Yarom, 2000). This hierarchical organization is well exemplified in the tonotopical representation of the auditory cortex, in which each part of the cortex selectively responds to a frequency (Lakatos et al., 2005) as well as within brain structures (Ding et al., 2016; Giraud & Poeppel, 2012). However, the oscillatory activity of large networks can modulate the intrinsic oscillator properties of smaller networks or individual neurons (Steriade, 2001; Csicsvari et al., 2003). While high-frequency oscillations usually are restricted to a small neuronal space, larger networks are recruited during slow oscillations (Steriade, 2001; Csicsvari, Jamieson, Wise, & Buzsáki, 2003). This pattern of cross-frequency oscillatory activity has been proposed to serve

General Discussion

as a communication mechanism to transfer information from local circuits to large-scale brain networks (Canolty & Knight, 2010). Nonetheless, it remains unclear how the hierarchical organization and the interplay between frequencies across networks are translated into a functional response to a particular stimulus.

Several experiments explored how the different frequency bands interact within and between networks with cross-frequency coupling (CFC) analyses. In particular, phase-amplitude CFC allows describing the statistical dependence between the phase of two brain rhythms and provides a plausible mechanism for linking activity that occurs at significantly different frequency rates (Canolty & Knight, 2010). These experiments have shown that the amplitude of high-frequency oscillations is generally modulated by low-frequency rhythms (Lakatos et al., 2005). This hierarchical organization of brain rhythms is proposed to subserve several cognitive mechanisms such as sensory processing (Händel & Haarmeier, 2009), attention (Schroeder & Lakatos, 2009), or memory (Tort et al., 2009). For instance, it has been shown that theta modulates gamma frequencies in the hippocampus (Bragin et al., 1995). In contrast, in the auditory cortex, delta modulates theta, which in turn modulates gamma (Giraud & Poeppel, 2012). Thus, high-frequency rhythms can be modulated by several low-frequency rhythms (Voytek et al., 2010). Some research has also shown that low-frequency activity can be modulated by rhythmic external sensory inputs (Lakatos et al., 2008; Luo & Poeppel, 2007; Saleh et al., 2010) as well as by internal cognitive processes such as learning and memory (Rizzuto & Pozzan, 2006).

Importantly, all the literature mentioned above explored the CFC of wide frequency bands (theta, gamma, delta). Here, we narrowed our investigation to the brain response to specific frequencies corresponding to the presented stimulus (words and syllables). This allowed us to assess the neuronal synchronization of brain networks to a specific stimulus frequency: the syllable and the word frequencies of the speech stream. In **Study II**, even though we did not directly assess CFC, we showed an increased connectivity strength (wPLI) at word and syllable frequencies between frontal and occipital regions in the learning compared to the random condition. This result suggests that changes in the connectivity patterns between large-scale networks could be related to word learning success. In this line, Sonty and colleagues (2007) showed that a reduction of the language network connectivity leads to poor performance in lexico-semantic judgment tasks in patients with aphasia. Notably, this reduction of connectivity occurred despite “apparently normal” patterns of activation in language-related regions. In the same direction, Kikuchi and colleagues (2011) found that an increased left parieto-temporal

General Discussion

coherence in the theta frequency range was correlated with a high performance of language-related tasks. These studies suggest that it is not the oscillatory activity *per se*, but it is the connectivity patterns that contribute to the language ability. Overall, in future studies would be interesting to further explore the role of connectivity patterns in language ability. Furthermore, it would be interesting to examine the interactions between local and long-distance brain regions at stimulus-specific frequencies (CFC, e.g., word and syllable frequencies).

Studies II and III also provided further evidence that when presenting a rhythmic stimulus, the brain response follows its rhythmic structure (Ding & Simon, 2012; Ding et al., 2016; Buiatti et al., 2009; Batterink & Paller, 2017). However, a remaining issue is which is its functional role, and which are the physiological bases. One of the most outstanding proposals to the functional reason for this phenomenon is to predict the inputs (Arnal et al., 2011). The predictive coding (PC) theory and other models of perception based upon hierarchical Bayesian inference (Mumford, 1992; Clark, 2013; Heilbron & Chait, 2018) suggest that the brain is constantly inferring the incoming inputs by constructing a hierarchical model about the world and this model is continuously updated to reduce the "prediction error" (Heilbron & Chait, 2018). The pattern regularities in the incoming sensory information are used to make those predictions (Barascud et al., 2015; Southwell et al., 2017; Chait et al., 2020). The information is hierarchically processed, with higher levels of processing modulating lower levels, e.g., by prioritizing specific inputs over others based on attention or previous experience. Some authors have shown that the selective extraction of relevant information occurs by aligning the brain activity to it, that is, by synchronizing the neuronal activity to the relevant stimuli rhythm (Schroeder & Lakatos, 2009). Thus, this mechanism would allow the brain to select and predict the incoming sensory information, which may facilitate information processing and provide a faster response. In line with this idea, in the learning condition of Study II, we showed an increased brain response to both words and syllables in the learning compared to the random condition. This finding, thus, supports the theory that better stimulus processing (and therefore better learning) occurs when the brain is synchronized with the external input.

Another important issue is the role of attention during stimulus processing. The PC theory, as well as other studies on oscillations, have shown that the neural synchronization to the stimulus increases when our attention is driven toward the stimulus, that is, when the stimulus is attended (Ding & Simon, 2012; Lakatos et al., 2008; 2013; Hillyard et al., 1998; Keitel et al., 2013; Ding et al., 2016). This body of literature evidenced that top-down attentional mechanisms can modulate brain oscillations. These studies open the window to study how oscillations and

General Discussion

neuronal synchronization with the stimulus can modify sensory perception. In **Study II**, we went one step forward by suggesting that attentional mechanisms may influence sensory processing and higher cognitive processes like language learning. This hypothesis is based on the enhanced long-distance connectivity strength found between fronto-occipital and fronto-parietal regions. We proposed that visual cues modulate visual attention, facilitating learning. The results of our experiment open the door to explore further relations between oscillations and other higher cognitive functions such as comprehension and general cognition.

6.6 Limitations and future directions

Although we hope that the results showed in this dissertation have contributed to expanding our previous knowledge on the mechanisms involved in language learning, some methodological considerations and limitations are worth considering for future research.

Firstly, the results of **Study I** must be carefully interpreted to avoid falling on the "reverse inference" problem. This concept refers to inferring specific cognitive processes from locations or patterns of neural activations found during experimental tasks (Nathan & Del Pinal, 2017). In *Experiment 1*, we reported a frontal distribution of the N400 during the initial stages of learning that shifted to parietal sites. We related the frontal topography with the involvement of domain-general processes such as attention and executive functions. This interpretation should be carefully treated as brain regions typically underly a variety of cognitive functions and, thus, we cannot directly relate the activations in certain brain regions with a specific cognitive function. Nonetheless, even though we cannot make strong statements, our suggestion is supported by vast literature relating frontal resources with domain-general cognitive mechanisms such as attention (Hopfinger, Buonocore & Magnun, 2000; Tzourio et al., 1997; see Stuss, 2005 for a review), executive functions (Stuss, 2011; Aron, 2008; Sauseng et al., 2005) and cognitive control (Badre et al., 2009; Bunge et al., 2002). Furthermore, despite the indirectness of the measure, the correlation found between the modulation of the N400 component across sessions and the verbal fluency task, a task previously associated with cognitive control (Hughes & Bryan, 2002; Carpenter et al., 2020), points in the same direction. Nonetheless, to clarify the underlying sources of the shift of topography, further source analyses or other experiments with neuroimaging techniques such as MEG should be performed.

Secondly, some of the results reported in the studies may be accounted for individual differences between participants. Those differences may originate from using different learning

General Discussion

strategies or cognitive abilities such as working memory or executive functions. In **Study II** and **III**, we used speech segmentation tasks based on SL. SL relies on an implicit learning mechanism assumed to occur in an incidental, automatic, non-conscious way (Schapiro & Turk-Browne, 2015). Instead, we used an associative word-learning task with explicit instructions to learn the associations in **Study I**. Thus, in the former, more conscious learning strategies may be used by participants to learn the associations. Participants may use different strategies such as self-generated cues (Freed & Marshall, 1995) or the phonological association of the novel word to similar existent words. Novel word-forms can activate existing lexical representations that share some phonological similarities with existing word-forms (Collin & Loftus, 1975). At the same time, novel word-forms are proposed to be integrated into the mental lexicon by creating semantical links with words already integrated. Therefore, one of the learners' plausible strategies to acquire novel words is to create a semantical or phonological relationship between the novel and existing word-forms (Perfetti, 2007). The different learning strategies used by learners in different tasks should be examined in more detail in future research.

Some limitations arise from the EEG techniques used in **Study I and II**. EEG has an excellent temporal resolution that allows the precise study of the brain's temporal dynamics. However, the spatial resolution is quite low and does not allow to extend of the findings to specific brain regions. Thus, future studies using techniques with higher spatial resolution such as fMRI or MEG should be performed to decipher the brain sources found in the present studies. Especially, it would be interesting to replicate **Study II** in an fMRI adapted experimental setting to explore the neural sources of fronto-occipital connectivity patterns found in the connectivity analysis.

In addition to methodological limitations, there were other limitations related to the experimental design. In both **Study I** and **III**, a larger sample size could provide clearer results. Especially in **Study III**, the small number of patients with electrodes implanted in MTL regions limited the sample size of the study. Some studies have shown that the anterior hippocampus seems to be more dominant in SL than the posterior hippocampus (Schapiro et al., 2012; 2017). A larger patient sample size with more diverse electrode locations may shed more light on specific activation patterns within different hippocampal regions during language learning. Furthermore, an increased sample size would allow exploring whether right and left brain structures show different activation patterns, as language processing has typically been related to the left hemisphere (Breitenstein et al., 2005) whereas music-related stimuli seem to entrain right hemisphere structures (Herholz et al., 2015; Sagi et al., 2012). It has also been shown that several brain areas sensitive to statistical speech structures, including temporal regions, are also

General Discussion

sensitive to non-speech sequences (Tremblay, Baroni & Hasson, 2013). Thus, it would be interesting to explore the neural response of temporal structures including the hippocampus and the auditory cortex in SL learning tasks when using tones instead of pseudo-words (Schön et al. 2008; François et al., 2014). In this sense, many studies have shown selectivity for speech and non-speech vocalizations of temporal regions (Norman-Haignere et al., 2015; 2019; Overath et al., 2015). Thus, future studies should further explore the role of stimulus type (Norman-Haignere et al., 2019) as well as the specific contribution of lateralization patterns on the hierarchical processing of the auditory sequences.

Another limitation of the studies carried out in this dissertation is that language learning paradigms were performed in controlled laboratory settings. These environments are inherently different from real-life situations, containing additional cues and rich environments (Jusczyk, 1999). In **Study II** and **III**, the speech segmentation paradigms are probably the ones that differ most from the natural environment because real languages do not contain only words of the same length. Instead, most real languages contain words of different lengths. Although infants easily segment trisyllabic words from the speech (Saffran et al., 1996), they fail when words have a different length (Johnson & Tyler, 2010) as occurs in natural speech. In future studies, it would be interesting to explore the neural entrainment to a more complex and naturalistic auditory stimulus (Ding & Simon, 2012; Luo & Poeppel, 2007; Pasley et al., 2012), for instance, including pseudo-words of different lengths to see whether similar results are found. The paradigms of associative word learning used in **Study I** also slightly differed from naturalistic situations. Usually, novel words appear embedded in contextual sentences or situations rather than in isolation (as occurs in our experiment). Despite the limitations, other studies using more naturalistic approaches have found similar results regarding the learning abilities and its neural correlates and the N400 modulation (Mestres-Missé et al., 2007; Borovsky et al., 2010; Frishkoff et al., 2010; Batterink & Neville 2011; Bakker et al., 2015; Mills et al., 2005).

6.7 Conclusions

In the present Ph.D. thesis, I combined different word learning paradigms with electrophysiological (EEG & intracortical EEG) and behavioral measures to explore language learning mechanisms in adults and children. Specifically, the thesis included three studies in which two aspects of word learning, the speech segmentation and the meaning mapping process, were explored. We developed three different paradigms in which participants could learn novel word forms from an artificial speech stream (Study II & III) or by associating the novel word forms to familiar (Study II) or unfamiliar (Study I) objects. The combination of several learning paradigms has allowed us to explore how multimodal learning occurs in adults and children and its neural bases. Based on the research aims delineated in chapter 2, the conclusions that can be drawn from the present doctoral thesis are the following:

- Word learning occurs at different time scales with a gradual initial encoding to a later word consolidation.
- The different time scales during word learning are accompanied by changes in the topographical distribution of the electrophysiological signal (N400), suggesting the involvement of different neuronal sources in both processes.
- Only consistent object-word associations facilitate the encoding of novel information
- Speech segmentation and meaning mapping can occur in parallel in children (as occurs in adults).
- The presence of multi-modal cues facilitates the word learning process in children.
- This multi-modal facilitation is driven by an increased phase synchronization between fronto-occipital regions.
- The brain synchronization during word learning might be driven by attentional mechanisms during the learning process.
- Not only language-specific brain structures are involved in word learning but also not language-specific structures such as the hippocampus are involved in the learning process.
- The word learning process seems to be driven by an interplay between language-specific and general domain resources.

General Discussion

Despite the insights presented here, the word learning process is very complex with many unresolved questions yet. I believe that in order to have a full understanding of the word learning process, it is indispensable to look with a multidisciplinary point of view. I hope the findings of this thesis may help to better understand the word acquisition process and pave the way for a new research. More specifically, I hope the present work may provide new insights for possible interventions for patients with language impairments such as aphasia or dyslexia as well as insights for new learning methods in language learning.

Chapter 7

References

Chapter 7 References

- Abla, D., Katahira, K., & Okanoya, K. (2008). On-line assessment of statistical learning by event-related potentials. *Journal of Cognitive Neuroscience*, 20(6), 952-964.
- Abla, D., & Okanoya, K. (2009). Visual statistical learning of shape sequences: An ERP study. *Neuroscience Research*, 64(2), 185-190.
- Abutalebi, J., & Green, D. (2007). Bilingual language production: The neurocognition of language representation and control. *Journal of neurolinguist.* 20(3), 242-275.
- Abutalebi, J. (2008). Neural aspects of second language representation and language control. *Acta Psychologica*. 128(3), 466-478.
- Abutalebi, J., Della Rosa, P., Ding, G., Weekes, B., Costa, A., & Green, D. (2013). Language proficiency modulates the engagement of cognitive control areas in multilinguals. *Cortex*. 49(3), 905-911.
- Ahissar, E., Nagarajan, S., Ahissar, M., Protopapas, A., Mahncke, H., & Merzenich, M. M. (2001). Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proceedings of the National Academy of Sciences*, 98(23), 13367-13372.
- Akhtar, N., Carpenter, M., & Tomasello, M. (1996). The role of discourse novelty in early word learning. *Child Development*, 67(2), 635-645.
- Alain, C., & Bernstein, L. J. (2008). From sounds to meaning: the role of attention during auditory scene analysis. *Current Opinion in Otolaryngology & Head and Neck Surgery*, 16(5), 485-489.
- Angwin, A. J., Phua, B., & Copland, D. A. (2014). Using semantics to enhance new word learning: An ERP investigation. *Neuropsychologia*. 59, 169-178.
- Ansaldo, U., Enfield, N.J. (2016). Editorial: Is the Language Faculty Nonlinguistic? *Frontiers in Psychology*. 7:861
- Arnal, L. H., Wyart, V., & Giraud, A. L. (2011). Transitions in neural oscillations reflect prediction errors generated in audiovisual speech. *Nature Neuroscience*, 14(6), 797.
- Aron, A. R. (2008). Progress in executive-function research: From tasks to functions to regions to networks. *Current Directions in Psychological Science*, 17(2), 124-129.
- Aslin, R. N., Pisoni, D. B., Hennessy, B. L., & Perey, A. J. (1981). Discrimination of voice onset time by human infants: New findings and implications for the effects of early experience. *Child Development*, 52(4), 1135.
- Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, 9(4), 321-324.
- Assaneo, M. F., & Poeppel, D. (2018). The coupling between auditory and motor cortices is rate-

References

- restricted: Evidence for an intrinsic speech-motor rhythm. *Science Advances*, 4(2), eaao3842.
- Astheimer, L. B., & Sanders, L. D. (2011). Predictability affects early perceptual processing of word onsets in continuous speech. *Neuropsychologia*, 49(12), 3512-3516.
- Axelsson, E. L., Churchley, K., & Horst, J. S. (2012). The right thing at the right time: Why ostensive naming facilitates word learning. *Frontiers in Psychology*, 3, 88.
- Baddeley, A. D., & Hitch, G. (1974). Working memory. In *Psychology of learning and motivation* (Vol. 8, pp. 47-89). Academic press.
- Baddeley, A., Papagno, C., & Vallar, G. (1988). When long-term learning depends on short-term storage. *Journal of Memory and Language*, 27(5), 586-595.
- Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychology Review*, 105(1), 158.
- Badre, D., & Wagner, A.D. (2002) Semantic retrieval, mnemonic control, and prefrontal Cortex. *Behavioral Cognitive Neuroscience Review*. 1, 206–18.
- Badre, D., Poldrack, R. A., Paré-Blagoev, E. J., Insler, R. Z., & Wagner, A. D. (2005). Dissociable controlled retrieval and generalized selection mechanisms in ventrolateral prefrontal cortex. *Neuron*, 47(6), 907-918.
- Badre, D., Hoffman, J., Cooney, J. W., & D'esposito, M. (2009). Hierarchical cognitive control deficits following damage to the human frontal lobe. *Nature Neuroscience*, 12(4), 515-522.
- Bailey, A., Phillips, W., & Rutter, M. (1996). Autism: towards an integration of clinical, genetic, neuropsychological, and neurobiological perspectives. *Journal of Child Psychology and Psychiatry*, 37(1), 89-126.
- Bakker, I., Takashima, A., van Hell, J. G., Janzen, G., & Mc Queen, J. M. (2015). Tracking lexical consolidation with ERPs: Lexical and semantic-priming effects on N400 and LPC responses to newly-learned words. *Neuropsychologia*, 79, 33-41.
- Baldassano, C., Chen, J., Zadbood, A., Pillow, J. W., Hasson, U., & Norman, K. A. (2017). Discovering event structure in continuous narrative perception and memory. *Neuron*, 95(3), 709-721.
- Baldwin, D. A. (1995). Understanding the link between joint attention and language. *Joint Attention: Its Origins and Role in Development*, 131, 158.
- Barascud, N., Aukstulewicz, R., Petsas, T., Zhao, S., Pearce, M., Friston, K., Chait, M. (2015). Sensitivity to the emergence of predictable structure in sound sequences. *Psychophysiology*. 52, s14.
- Barrett, S. E., & Rugg, M. D. (1990). Event-related potentials and the semantic matching of pictures. *Brain and Cognition*, 14(2), 201-212.
- Barrett, M., *The development of Language*. (1999). United Kingdom, Sussex. *Psychology Press*

References

- Bartolotti, J., Bradley, K., Hernandez, A. E. & Marian, V. (2017). Neural signatures of second language learning and control. *Neuropsychologia*, 98, 130-138.
- Bastos, A. M., & Schoffelen, J. M. (2016). A tutorial review of functional connectivity analysis methods and their interpretational pitfalls. *Frontiers in Systems Neuroscience*, 9, 175.
- Batterink, L., & Neville, H. (2011). Implicit and explicit mechanisms of word learning in a narrative context: an event-related potential study. *Journal Cognitive Neuroscience*. 23(11), 3181-3196.
- Batterink, L. J., Reber, P. J., Neville, H. J., & Paller, K. A. (2015). Implicit and explicit contributions to statistical learning. *Journal of Memory and Language*, 83, 62-78.
- Batterink, L. J., & Paller, K. A. (2017). Online neural monitoring of statistical learning. *Cortex*, 90, 31-45.
- Ben-Yakov, A., & Dudai, Y. (2011). Constructing realistic engrams: poststimulus activity of hippocampus and dorsal striatum predicts subsequent episodic memory. *Journal of Neuroscience*, 31(24), 9032-9042.
- Bermúdez-Margaretto B., Beltrán D., Domínguez A., & Cuetos F. (2015). Repeated exposure to 'meaningless' pseudowords modulates LPC, but not N(FN)400. *Brain Topography*. 28, 838-51.
- Bermúdez-Margaretto, B., Beltrán, D., Cuetos, F., & Domínguez, A. (2018). Brain signatures of new (Pseudo-) words: visual repetition in associative and non-associative contexts. *Frontiers in Human Neuroscience*. 12, 354.
- Bermúdez-Margaretto, B., Beltrán, D., Cuetos, F., & Domínguez, A. (2019). Novel word learning: event-related brain potentials reflect pure lexical and task-related effects. *Frontiers in Human Neuroscience*. 13, 347.
- Birch, S. A., Vauthier, S. A., & Bloom, P. (2008). Three- and four-year-olds spontaneously use others' past performance to guide their learning. *Cognition*, 107(3), 1018-1034.
- Bischoff-Grethe, A., Proper, S. M., Mao, H., Daniels, K. A., & Berns, G. S. (2000). Conscious and unconscious processing of nonverbal predictability in Wernicke's area. *Journal of Neuroscience*, 20(5), 1975-1981.
- Bloom, P. (1998). Language acquisition in its developmental context.
- Bloom P. (2000). How children learn the meanings of words. (Cambridge: MIT Press).
- Borgström, K., von Koss Torkildsen, J., & Lindgren, M. (2015). Substantial gains in word learning ability between 20 and 24 months: A longitudinal ERP study. *Brain and Language*, 149, 33-45.
- Borovsky, A., Kutas, M., & Elman, J. (2010). Learning to use words: Event-related potentials index single-shot contextual word learning. *Cognition*, 116(2), 289-296.
- Bradley, K. A., King, K. E., & Hernandez, A. E. (2013). Language experience differentiates prefrontal and subcortical activation of the cognitive control network in novel word

References

- learning. *NeuroImage*. 67, 101-110.
- Bragin, A., Jandó, G., Nádasdy, Z., Hetke, J., Wise, K., & Buzsáki, G. (1995). Gamma (40-100 Hz) oscillation in the hippocampus of the behaving rat. *Journal of Neuroscience*, 15(1), 47-60.
- Breitenstein, C., Jansen, A., Deppe, M., Foerster, A. F., Sommer, J., Wolbers, T., & Knecht, S. (2005). Hippocampus activity differentiates good from poor learners of a novel lexicon. *Neuroimage*, 25(3), 958-968.
- Briellmann, R. S., Saling, M. M., Connell, A. B., Waites, A. B., Abbott, D. F., & Jackson, G. D. (2004). A high-field functional MRI study of quadri-lingual subjects. *Brain & Language*. 89(3), 531-542.
- 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.
- Brod S. (2018). Fast track to the neocortex: A memory engram in the posterior parietal cortex. *Science*. 362, 1045-1048.
- Buiatti, M., Peña, M., & Dehaene-Lambertz, G. (2009). Investigating the neural correlates of continuous speech computation with frequency-tagged neuroelectric responses. *Neuroimage*, 44(2), 509-519.
- Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., & Gabrieli, J. D. (2002). Immature frontal lobe contributions to cognitive control in children: evidence from fMRI. *Neuron*, 33(2), 301-311.
- Buzsáki, G. & Dragun A. (2004). Neuronal oscillations in cortical networks. *Science*, 304, 1926-1929
- Buzsáki, G., & Wang, X. J. (2012). Mechanisms of gamma oscillations. *Annual review of Neuroscience*, 35, 203-225.
- Campbell, K. L., & Tyler, L. K. (2018). Language-related domain-specific and domain-general systems in the human brain. *Current Opinion in Behavioral Sciences*, 21, 132-137.
- Canolty, R. T., Edwards, E., Dalal, S. S., Soltani, M., Nagarajan, S. S., Kirsch, H. E., M.S. Berger, Barbaro, N. M., & Knight, R. T. (2006). High gamma power is phase-locked to theta oscillations in human neocortex. *Science*, 313(5793), 1626-1628.
- Canolty, R. T., & Knight, R. T. (2010). The functional role of cross-frequency coupling. *Trends in Cognitive Sciences*, 14(11), 506-515.
- Capilla, A., Pazo-Alvarez, P., Darriba, A., Campo, P., & Gross, J. (2011). Steady-state visual evoked potentials can be explained by temporal superposition of transient event-related responses. *PLoS One*, 6(1), e14543.
- Carey S., Bartlett E. (1978) Acquiring a single new word. *Papers and Reports on Child Language Development*. 15, 17-29.
- Carpenter, E., Rao, L., Peñaloza, C., & Kiran, S. (2020). Verbal fluency as a measure of lexical

References

- access and cognitive control in bilingual persons with aphasia. *Aphasiology*, 1-22.
- Catani, M., Jones, D. K., & Ffytche, D. H. (2005). Perisylvian language networks of the human brain. *Annals of Neurology: Official Journal of the American Neurological Association and the Child Neurology Society*, 57(1), 8-16.
- Catani, M., Allin, M. P., Husain, M., Pugliese, L., Mesulam, M. M., Murray, R. M., & Jones, D. K. (2007). Symmetries in human brain language pathways correlate with verbal recall. *Proceedings of the National Academy of Sciences*, 104(43), 17163-17168.
- Catani, M., & De Schotten, M. T. (2008). A diffusion tensor imaging tractography atlas for virtual in vivo dissections. *Cortex*, 44(8), 1105-1132.
- Chait, M., de Cheveigné, A., Poeppel, D., & Simon, J. Z. (2010). Neural dynamics of attending and ignoring in human auditory cortex. *Neuropsychologia*, 48(11), 3262-3271.
- Chait, M. (2020). How the brain discovers structure in sound sequences. *Acoustical Science and Technology*, 41(1), 48-53.
- Chaturvedi, M., Bogaarts, J. G., Kozak, V. V., Hatz, F., Gschwandtner, U., Meyer, A., ... & Roth, V. (2019). Phase lag index and spectral power as QEEG features for identification of patients with mild cognitive impairment in Parkinson's disease. *Clinical Neurophysiology*, 130(10), 1937-1944.
- Chein, J. M., & Schneider, W. (2005). Neuroimaging studies of practice-related change: fMRI and meta-analytic evidence of a domain-general control network for learning. *Cognitive Brain Research*. 25(3), 607-623.
- Chein, J. M., & Schneider, W. (2012). The brain's learning and control architecture. *Current Directions in Psychological Science*. 21(2), 78-84.
- Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the acoustical society of America*, 25(5), 975-979.
- Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.
- Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. Greenwood Publishing Group.
- Clark, E. V., & MacWhinney, B. (1987). The principle of contrast: A constraint on language acquisition. *Mechanisms of language acquisition*, 1-33.
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and brain sciences*, 36(3), 181-204.
- Clarke, S., Thiran, A. B., Maeder, P., Adriani, M., Vernet, O., Regli, L., Cuisenaire, O., & Thiran, J. P. (2002). What and where in human audition: selective deficits following focal hemispheric lesions. *Experimental Brain Research*, 147(1), 8-15.
- Clements-Stephens, A. M., Materek, A. D., Eason, S. H., Scarborough, H. S., Pugh, K. R., Rimrodt, S., & Cutting, L. E. (2012). Neural circuitry associated with two different approaches to novel word learning. *Developmental Cognitive Neuroscience*. 2, S99-S113.

References

- Cohen, N. J., & Eichenbaum, H. B. (1993). *Memory, amnesia, and hippocampal function*. Cambridge: MIT Press.
- Colombo, P. J., Davis, H. P., & Volpe, B. T. (1989). Allocentric spatial and tactile memory impairments in rats with dorsal caudate lesions are affected by preoperative behavioral training. *Behavioral Neuroscience*, 103, 1242-1250.
- Cohen M.X. *Analyzing Neural Time Series Data: Theory and Practice*. (2014). London, England. The MIT Press.
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82(6), 407.
- Colombet, B., Woodman, M., Badier, J. M. & Bénar, C. G. (2015). AnyWave: A cross-platform and modular software for visualizing and processing electrophysiological signals. *Journal of Neuroscience Methods*, 242, 118-126.
- Coran, M., Rodríguez-Fornells, A., Ramos-Escobar, N., Laine, M., & Martin, N. (2020). Word learning in aphasia: Treatment implications and structural connectivity analyses. *Topics in Language Disorders*, 40(1), 81-109.
- Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience*, 3(3), 201-215.
- Coutanche, M. N., & Thompson-Schill, S. L. (2014). Fast mapping rapidly integrates information into existing memory networks. *Journal of Experimental Psychology*. General 143(6), 2296-2303.
- Covic, A., Keitel, C., Porcu, E., Schröger, E., & Müller, M. M. (2017). Audio-visual synchrony and spatial attention enhance processing of dynamic visual stimulation independently and in parallel: a frequency-tagging study. *Neuroimage*, 161, 32-42.
- Covington, N. V., Duff, M. C. (2016). Expanding the language network: Direct contributions from the hippocampus. *Trends in Cognitive Sciences*, 20(12), 869-870.
- Covington, N. V., Brown-Schmidt, S., & Duff, M. C. (2018). The necessity of the hippocampus for statistical learning. *Journal of Cognitive Neuroscience*, 30(5), 680-697.
- Csicsvari, J., Jamieson, B., Wise, K. D., & Buzsáki, G. (2003). Mechanisms of gamma oscillations in the hippocampus of the behaving rat. *Neuron*, 37(2), 311-322.
- Cunillera, T., Toro, J. M., Sebastián-Gallés, N., & Rodríguez-Fornells, A. (2006). The effects of stress and statistical cues on continuous speech segmentation: an event-related brain potential study. *Brain Research*, 1123(1), 168-178.
- Cunillera, T., Càmara, E., Toro, J. M., Marco-Pallares, J., Sebastián-Gallés, N., Ortiz, H., Pujol, J., & Rodríguez-Fornells, A. (2009). Time course and functional neuroanatomy of speech segmentation in adults. *Neuroimage*, 48(3), 541-553.
- Cunillera, T., Laine, M., Càmara, E., & Rodríguez-Fornells, A. (2010a). Bridging the gap between speech segmentation and word-to-world mappings: Evidence from an audio-visual statistical learning task. *Journal of Memory and Language*, 63(3), 295-305.
- Cunillera, T., Càmara, E., Laine, M., & Rodríguez-Fornells, A. (2010b). Speech segmentation is

References

- facilitated by visual cues. *Quarterly Journal of Experimental Psychology*, 63(2), 260-274.
- Cunillera, T., Fuentemilla, L., Periañez, J., Marco-Pallarès, J., Krämer, U. M., Càmara, E., Münte T.F., Rodríguez-Fornells, A. (2012). Brain oscillatory activity associated with task switching and feedback processing. *Cognitive, Affective, & Behavioral Neuroscience*, 12(1), 16-33.
- Curran, T. (2000). Brain potentials of recollection and familiarity. *Memory & cognition*, 28(6), 923-938.
- Cutler, A., & Carter, D. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer speech and Language*, 2, 133-142.
- Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, 14(1), 113.
- Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language*, 31(2), 218-236.
- Davachi, L. (2006). Item, context and relational episodic encoding in humans. *Current Opinion in Neurobiology*, 16(6), 693-700.
- Davis, C. J., Perea, M. A. (2005) Program for deriving orthographic and phonological neighborhood statistics and other psycholinguistic indices in Spanish. *Behavioral Research Methods*. 37, 665–671.
- Davis, M. H., & Gaskell, M. G. (2009). A complementary systems account of word learning: neural and behavioural evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1536), 3773-3800.
- De Diego-Balaguer, R. D. D., Toro, J. M., Rodríguez-Fornells, A., & Bachoud-Lévi, A. C. (2007). Different neurophysiological mechanisms underlying word and rule extraction from speech. *PLoS One*, 2(11), e1175.
- de Diego-Balaguer, R., Martínez-Alvarez, A., & Pons, F. (2016). Temporal attention as a scaffold for language development. *Frontiers in Psychology*, 7, 44.
- Deacon, D., Dynowska, A., Ritter, W. & Grose-Fifer, J. (2004). Repetition and semantic priming of nonwords: implications for theories of N400 and word recognition. *Psychophysiology*. 41, 60–74.
- Demb, J.B., Desmond, J.E., Wagner, A.D., Vaidya, C.J., Glover, G.H. & Gabrieli, J.D. (1995). Semantic encoding and retrieval in the left Inferior Prefrontal Cortex: A functional MRI study of task difficulty and process specificity. *Journal of Neuroscience*. 15, 5870–5878.
- Destrieux, C., Fischl, B., Dale, A. & Halgren, E. (2010). Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuroimage*, 53(1), 1-15.
- Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy, and biochemistry. *Principles of Frontal Lobe Function*, 466-503.
- Dick, A. S., & Tremblay, P. (2012). Beyond the arcuate fasciculus: consensus and controversy in

References

- the connective anatomy of language. *Brain*, 135(12), 3529-3550.
- Diehl, B., Busch, R. M., Duncan, J. S., Piao, Z., Tkach, J., & Lüders, H. O. (2008). Abnormalities in diffusion tensor imaging of the uncinate fasciculus relate to reduced memory in temporal lobe epilepsy. *Epilepsia*, 49(8), 1409-1418.
- Ding, N., Melloni, L., Zhang, H., Tian, X., & Poeppel, D. (2016). Cortical tracking of hierarchical linguistic structures in connected speech. *Nature Neuroscience*, 19(1), 158-164.
- Dittinger, E., Barbaroux, M., D'Imperio, M., Jäncke, L., Elmer, S., & Besson, M. (2016). Professional music training and novel word learning: From faster semantic encoding to longer-lasting word representations. *Journal of Cognitive Neuroscience*, 28, 1584–1602.
- Dittinger, E., Chobert, J., Ziegler, J. C., & Besson, M. (2017). Fast brain plasticity during word learning in musically-trained children. *Frontiers in Human Neuroscience*, 11, 233.
- Dobel, C., Lagemann, L., & Zwitserlood, P. (2009). Non-native phonemes in adult word learning: evidence from the N400m. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1536), 3697-3709.
- Doelling, K. B., Arnal, L. H., Ghizva, O., & Poeppel, D. (2014). Acoustic landmarks drive delta–theta oscillations to enable speech comprehension by facilitating perceptual parsing. *Neuroimage*, 85, 761-768.
- Doelling, K. B., Assaneo, M. F., Bevilacqua, D., Pesaran, B., & Poeppel, D. (2019). An oscillator model better predicts cortical entrainment to music. *Proceedings of the National Academy of Sciences*, 116(20), 10113-10121.
- Doesburg, S. M., Vinette, S. A., Cheung, M. J., & Pang, E. W. (2012). Theta-modulated gamma-band synchronization among activated regions during a verb generation task. *Frontiers in Psychology*, 3, 195.
- Duchaine, B. C., & Nakayama, K. (2006). Developmental prosopagnosia: a window to content-specific face processing. *Current Opinion in Neurobiology*, 16(2), 166-173.
- Duff, M. C., & Brown-Schmidt, S. (2012). The hippocampus and the flexible use and processing of language. *Frontiers in Human Neuroscience*, 6, 69.
- Dumay, N., & Gaskell, M.G. (2012). Overnight lexical consolidation revealed by speech segmentation. *Cognition*, 123, 119-132.
- Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, 171(3968), 303-306.
- Elgort, I., Perfetti, C. A., Rickles, B., & Stafura, J. Z. (2015). Contextual learning of L2 word meanings: Second language proficiency modulates behavioural and event-related brain potential (ERP) indicators of learning. *Language Cognition & Neuroscience*, 30(5), 506-528.
- Elmer, S., Abolfazl, S., Cunillera T., Rodriguez-Fornells, A., (in prep). Selective Neural Entrainment to Speech Units While Learning New Words Based on Statistical Regularities and Prosodic Cues.

References

- Embury, C. M., Wiesman, A. I., Proskovec, A. L., Mills, M. S., Heinrichs-Graham, E., Wang, Y. P., Calhoun V.D., Stephen M.J., & Wilson, T. W. (2019). Neural dynamics of verbal working memory processing in children and adolescents. *NeuroImage*, 185, 191-197.
- Fadiga, L., Craighero, L., Buccino, G., & Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *European journal of Neuroscience*, 15(2), 399-402.
- Farthouat, J., Franco, A., Mary, A., Delpouve, J., Wens, V., De Beeck, M. O., De Tiège, X., & Peigneux, P. (2017). Auditory magnetoencephalographic frequency-tagged responses mirror the ongoing segmentation processes underlying statistical learning. *Brain Topography*, 30(2), 220-232.
- Federmeier, K.D., McLennan, D.B., De Ochoa, E., & Kutas, M. (2002). The impact of semantic memory organization and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology*, 39(2), 133-146.
- Federmeier, K.D., Kutas, M. & Schul, R. (2010). Age-related and individual differences in the use of prediction during language comprehension. *Brain & Language*, 115, 149-61.
- Fedorenko, E. (2014). The role of domain-general cognitive control in language comprehension. *Frontiers in Psychology*, 5, 335.
- Ferreira, R. A., Göbel, S. M., Hymers, M., & Ellis, A. W. (2015). The neural correlates of semantic richness: Evidence from an fMRI study of word learning. *Brain & Language*. 143, 69-80.
- Ferreira, R., & Bernales, C. (2020). Exploring foreign language learning with EEG and eye-tracking: what can online tools add to the story? *English as a Foreign Language: Perspectives on Teaching, Multilingualism and Interculturalism*. Cambridge Scholars Publishing.
- Ferstl, E., Neumann, J., Bogler, C., & Von Cramon, D.Y. (2008). The extended language network: A meta-analysis of neuroimaging studies on text comprehension. *Human Brain Mapping*. 29, 581-593.
- Fiser, J., & Aslin, R. N. (2002a). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(3), 458.
- Fiser, J., & Aslin, R. N. (2002b). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences*, 99(24), 15822-15826.
- Fitch, W. T. (2010). *The evolution of language*. Cambridge University Press.
- Fletcher, P. C., Zafiris, O., Frith, C. D., Honey, R. A. E., Corlett, P. R., Zilles, K., & Fink, G. R. (2005). On the benefits of not trying: Brain activity and connectivity reflecting the interactions of explicit and implicit sequence learning. *Cerebral Cortex*, 15(7), 1002-1015.
- Francois, C., & Schön, D. (2011). Musical expertise and statistical learning of musical and linguistic structures. *Frontiers in Psychology*, 2, 167.
- Francois, C., Tillmann, B., & Schon, D. (2012). Cognitive and methodological considerations on

References

- the effects of musical expertise on speech segmentation. *Annals of the New York Academy of Sciences*, 108-115.
- François, C., Chobert, J., Besson, M., & Schön, D. (2013). Music training for the development of speech segmentation. *Cerebral Cortex*, 23(9), 2038-2043.
- François, C., Jaillet, F., Takerkart, S., & Schön, D. (2014). Faster sound stream segmentation in musicians than in nonmusicians. *PLoS One*, 9(7), e101340.
- François, C., Ripollés, P., Bosch, L., Garcia-Alix, A., Muchart, J., Sierpowska, J., Fons, C., Solé, J., Rebollo, M., Gaitán, H., & Rodriguez-Fornells, A. (2016). Language learning and brain reorganization in a 3.5-year-old child with left perinatal stroke revealed using structural and functional connectivity. *Cortex*, 77, 95-118.
- François, C., Teixidó, M., Takerkart, S., Agut, T., Bosch, L., & Rodriguez-Fornells, A. (2017a). Enhanced neonatal brain responses to sung streams predict vocabulary outcomes by age 18 months. *Scientific Reports*, 7(1), 1-13.
- François, C., Cunillera, T., Garcia, E., Laine, M., & Rodriguez-Fornells, A. (2017b). Neurophysiological evidence for the interplay of speech segmentation and word-referent mapping during novel word learning. *Neuropsychologia*, 98, 56-67.
- Freed, D. B., & Marshall, R. C. (1995). The effect of personalized cueing on long-term naming of realistic visual stimuli. *American Journal of Speech-Language Pathology*, 4(4), 105-108.
- Frey, S., Kostopoulos, P., & Petrides, M. (2004). Orbitofrontal contribution to auditory encoding. *Neuroimage*, 22(3), 1384-1389.
- Friedrich, M., & Friederici, A. D. (2004). N400-like semantic incongruity effect in 19-month-olds: Processing known words in picture contexts. *Journal of Cognitive Neuroscience*, 16(8), 1465-1477.
- Friedrich, M., & Friederici, A. D. (2008). Neurophysiological correlates of online word learning in 14-month-old infants. *Neuroreport*, 19(18), 1757-1761.
- Friedrich, M., & Friederici, A. D. (2011). Word learning in 6-month-olds: fast encoding–weak retention. *Journal of Cognitive Neuroscience*, 23(11), 3228-3240.
- Friedrich, M., Friederici, A.D. (2017). The origins of word learning: Brain responses of 3-month-olds indicate their rapid association of objects and words. *Developmental Science*. 20, 12357 (2017).
- Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, 9(10), 474-480.
- Frishkoff, G.A., Perfetti, C.A., Collins-Thompson, K. (2010). Lexical quality in the brain: ERP evidence for robust word learning from context. *Developmental Neuropsychology* 35, 376–403.
- Fritz, J., Mishkin, M., & Saunders, R. C. (2005). In search of an auditory engram. *Proceedings of the National Academy of Sciences*, 102(26), 9359-9364.

References

- Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: the paradox of statistical learning. *Trends in Cognitive Sciences*, 19(3), 117-125.
- Ganis, G., Kutas, M., & Sereno, M. I. (1996). The search for "common sense": An electrophysiological study of the comprehension of words and pictures in reading. *Journal of Cognitive Neuroscience*, 8(2), 89-106.
- Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language*, 28(2), 200-213.
- Gathercole, S. E., & Baddeley, A. D. (1990). The role of phonological memory in vocabulary acquisition: A study of young children learning new names. *British Journal of Psychology*, 81(4), 439-454.
- Gathercole, S. E., Willis, C., Emslie, H., & Baddeley, A. D. (1991). The influences of number of syllables and wordlikeness on children's repetition of nonwords. *Applied Psycholinguistics*, 12(3), 349-367.
- Gathercole, S. E., & Baddeley, A. D. (1993). Phonological working memory: A critical building block for reading development and vocabulary acquisition?. *European Journal of Psychology of Education*, 8(3), 259.
- Gazzaniga, M. S. (Ed.). (2014). *Handbook of cognitive neuroscience*. Springer.
- Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition*, 23(3), 183-209.
- Ghitza, O., Giraud, A. L., & Poeppel, D. (2013). Neuronal oscillations and speech perception: critical-band temporal envelopes are the essence. *Frontiers in Human Neuroscience*, 6, 340.
- Gibson, J. J., & Gibson, E. J. (1955). Perceptual learning: Differentiation or enrichment?. *Psychological Review*, 62(1), 32.
- Gillette, J., Gleitman, H., Gleitman, L., Lederer, A., 1999. Human simulations of vocabulary learning. *Cognition* 73, 135-176.
- Giner, G. (2011). *El Jinete del silencio*. Barcelona. Editorial Planeta.
- Giraud, A. L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature Neuroscience*, 15(4), 511.
- Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition*, 1(1), 3-55.
- Glicksohn, A., & Cohen, A. (2013). The role of cross-modal associations in statistical learning. *Psychonomic Bulletin & Review*, 20(6), 1161-1169.
- Gold, B.T., Balota, D.A., Jones, S.J., Powell, D.K., Smith, C.D., Andersen, A.H. (2006). Dissociation of automatic and strategic lexical-semantics: Functional magnetic resonance imaging evidence for differing roles of multiple frontotemporal regions. *Journal of Neuroscience*. 26, 6523-6532.

References

- Gorno-Tempini, M. L., Murray, R. C., Rankin, K. P., Weiner, M. W., & Miller, B. L. (2004). Clinical, cognitive and anatomical evolution from nonfluent progressive aphasia to corticobasal syndrome: a case report. *Neurocase*, 10(6), 426-436.
- Graf-Estes, K., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science*, 18(3), 254-260.
- Greenberg, S., & Arai, T. (2004). What are the essential cues for understanding spoken language?. *IEICE TRANSACTIONS on Information and Systems*, 87(5), 1059-1070.
- Groppe, D. M., Bickel, S., Dykstra, A. R., Wang, X., Mégevand, P., Mercier, M. R., Lado, F. A., Mehta, A. D., & Honey, C. J. (2017). iELVis: An open source MATLAB toolbox for localizing and visualizing human intracranial electrode data. *Journal of neuroscience methods*. 281, 40-48.
- Gupta, P. (1996). Verbal short-term memory and language processing: A computational model. *Brain and Language*, 55, 194-97.
- Gupta, P. (2003). Examining the relationship between word learning, nonword repetition, and immediate serial recall in adults. *The Quarterly Journal of Experimental Psychology Section A*, 56(7), 1213-1236.
- Gupta, P., & Tisdale, J. (2009). Word learning, phonological short-term memory, phonotactic probability and long-term memory: towards an integrated framework. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1536), 3755-3771.
- Haber, S. N., & Calzavara, R. (2009). The cortico-basal ganglia integrative network: the role of the thalamus. *Brain Research Bulletin*, 78(2-3), 69-74.
- Hagoort, P. (2019). The neurobiology of language beyond single-word processing. *Science*. 366(6461), 55-58.
- Hämäläinen, M., Hari, R., Ilmoniemi, R. J., Knuutila, J., & Lounasmaa, O. V. (1993). Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics*, 65(2), 413.
- Hamm, J. P., Johnson, B. W., & Kirk, I. J. (2002). Comparison of the N300 and N400 ERPs to picture stimuli in congruent and incongruent contexts. *Clinical Neurophysiology*, 113(8), 1339-1350.
- Händel, B., & Haarmeier, T. (2009). Cross-frequency coupling of brain oscillations indicates the success in visual motion discrimination. *Neuroimage*, 45(3), 1040-1046.
- Hardmeier, M., Hatz, F., Bousleiman, H., Schindler, C., Stam, C. J., & Fuhr, P. (2014). Reproducibility of functional connectivity and graph measures based on the phase lag index (PLI) and weighted phase lag index (wPLI) derived from high resolution EEG. *PLoS One*, 9(10), e108648.
- Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve?. *Science*, 298(5598), 1569-1579.

References

- Havas, V., Taylor, J.S.H., Vaquero, L., de Diego-Balaguer, R., Rodríguez-Fornells, A., Davis, M.H. (2018) Semantic and phonological schema influence spoken word learning and overnight consolidation. *Quarterly Journal of Experimental Psychology*, 71, 1469-1481.
- Hay, J. F., Pelucchi, B., Estes, K. G., & Saffran, J. R. (2011). Linking sounds to meanings: Infant statistical learning in a natural language. *Cognitive Psychology*, 63(2), 93-106.
- Hebscher, M., Wing, E., Ryan, J., & Gilboa, A. (2019). Rapid Cortical Plasticity Supports Long-Term Memory Formation. *Trends Cognitive Science*, 23(12), 989-1002.
- Heilbron, M., & Chait, M. (2018). Great expectations: is there evidence for predictive coding in auditory cortex?. *Neuroscience*, 389, 54-73.
- Helfrich, R. F., Fiebelkorn, I. C., Szczepanski, S. M., Lin, J. J., Parvizi, J., Knight, R. T., & Kastner, S. (2018). Neural mechanisms of sustained attention are rhythmic. *Neuron*, 99(4), 854-865.
- Henin, S., Turk-Browne, N., Friedman, D., Liu, A., Dugan, P., Flinker, A., Doyle W., Devinsky, O., & Melloni, L. (2020). Statistical learning shapes neural sequence representations. *BioRxiv*, 583856.
- Herholz, S. C., Coffey, E. B., Pantev, C., & Zatorre, R. J. (2015). Dissociation of neural networks for predisposition and for training-related plasticity in auditory-motor learning. *Cerebral Cortex*, 26(7), 3125-3134.
- Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends in Cognitive Sciences*, 4(4), 131-138.
- Hickok, G., & Poeppel, D. (2004). Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*, 92(1-2), 67-99.
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8(5), 393-402
- Hill, M., & Wagovich, S. (2020). Word learning from context in school-age children: relations with language ability and executive function. *Journal of Child Language*, 1-24.
- Hillyard, S. A., Vogel, E. K., & Luck, S. J. (1998). Sensory gain control (amplification) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 353(1373), 1257-1270.
- Holcomb, P. J., & Mcpherson, W. B. (1994). Event-related brain potentials reflect semantic priming in an object decision task. *Brain and Cognition*, 24(2), 259-276.
- Hollich, G., Hirsh-Pasek, K., & Golinkoff, R. M. (2000). I. What Does it Take to Learn a Word? *Monographs of the Society for Research in Child Development*, 65(3), 1-16.
- Hopfinger, J. B., Buonocore, M. H., & Mangun, G. R. (2000). The neural mechanisms of top-down attentional control. *Nature Neuroscience*, 3, 284-291
- Horst, J. S., & Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy*, 13(2), 128-157.

References

- Hosoda, C., Tanaka, K., Nariai, T., Honda, M., & Hanakawa, T. (2013). Dynamic neural network reorganization associated with second language vocabulary acquisition: A multimodal imaging study. *Journal of Neuroscience*, 33(34), 13663-13672.
- Hughes, D. L., & Bryan, J. (2002). Adult age differences in strategy use during verbal fluency performance. *Journal of Clinical Experimental Neuropsychology*, 24(5), 642-654.
- Hulme, C., Maughan, S., & Brown, G. D. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language*, 30(6), 685-701.
- Hunt, R. H., & Aslin, R. N. (2001). Statistical learning in a serial reaction time task: access to separable statistical cues by individual learners. *Journal of Experimental Psychology: General*, 130(4), 658.
- Hutcheon, B., & Yarom, Y. (2000). Resonance, oscillation and the intrinsic frequency preferences of neurons. *Trends in Neurosciences*, 23(5), 216-222.
- James, W. (2007). *The principles of psychology* (Vol. 1). United States, New York. Cosimo, Inc.
- Jeon, H. A., & Friederici, A. D. (2015). Degree of automaticity and the prefrontal cortex. *Trends Cognitive Science*, 19(5), 244-250.
- Johnson, E. K., & Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language*, 44(4), 548-567.
- Johnson, E. K., & Tyler, M. D. (2010). Testing the limits of statistical learning for word segmentation. *Developmental Science*, 13(2), 339-345.
- Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21(1), 60-99.
- Junge, C., Kooijman, V., Hagoort, P., & Cutler, A. (2012). Rapid recognition at 10 months as a predictor of language development. *Developmental Science*, 15(4), 463-473.
- Jusczyk, P. W. (1999). How infants begin to extract words from speech. *Trends in Cognitive Sciences*, 3(9), 323-328.
- Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology*, 39(3-4), 159-207.
- Kabdebon, C., Pena, M., Buiatti, M., & Dehaene-Lambertz, G. (2015). Electrophysiological evidence of statistical learning of long-distance dependencies in 8-month-old preterm and full-term infants. *Brain and Language*, 148, 25-36.
- Kaczer, L., Bavassi, L., Petroni, A., Fernández, R. S., Laurino, J., Degiorgi, S., & Pedreira, M. E. (2018). Contrasting dynamics of memory consolidation for novel word forms and meanings revealed by behavioral and neurophysiological markers. *Neuropsychologia*, 117, 472-482.
- Kapa, L.L., & Colombo, J. (2014). Executive function predicts artificial language learning. *Journal of Memory and Language*, 76, 237-252.

References

- Karuza, E. A., Newport, E. L., Aslin, R. N., Starling, S. J., Tivarus, M. E., & Bavelier, D. (2013). The neural correlates of statistical learning in a word segmentation task: An fMRI study. *Brain and Language*, 127(1), 46-54.
- Kaya, E. M., & Elhilali, M. (2017). Modelling auditory attention. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1714), 20160101.
- Kayser, J., Tenke, C.E. (2006). Principal components analysis of Laplacian waveforms as a generic method for identifying ERP generator patterns: I. Evaluation with auditory oddball tasks. *Clinical Neurophysiology*, 117(2), 348-368.
- Keitel, C., Andersen, S. K., Quigley, C., & Müller, M. M. (2013). Independent effects of attentional gain control and competitive interactions on visual stimulus processing. *Cerebral Cortex*, 23(4), 940-946.
- Keitel, C., Quigley, C., & Ruhnau, P. (2014). Stimulus-driven brain oscillations in the alpha range: entrainment of intrinsic rhythms or frequency-following response? *Journal of Neuroscience*, 34(31), 10137-10140.
- Kemler Nelson, D. G., Russell, R., Duke, N., & Jones, K. (2000). Two-year-olds will name artifacts by their functions. *Child Development*, 71(5), 1271-1288.
- Keren-Happuch, E., Chen, S. H. A., Ho, M. H. R., & Desmond, J. E. (2014). A meta-analysis of cerebellar contributions to higher cognition from PET and fMRI studies. *Human Brain Mapping*, 35(2), 593.
- Kikuchi, M., Shitamichi, K., Yoshimura, Y., Ueno, S., Remijn, G. B., Hirosawa, T., Munesue, T., Tsubokawa, T., Haruta, Y., Oi, M., Higashida, H., & Minabe, Y. (2011). Lateralized theta wave connectivity and language performance in 2-to 5-year-old children. *Journal of Neuroscience*, 31(42), 14984-14988.
- Kim, A. S., Vallesi, A., Picton, T. W., & Tulving, E. (2009). Cognitive association formation in episodic memory: Evidence from event-related potentials. *Neuropsychologia*, 47(14), 3162-3173.
- Kim, Y. J., Grabowecky, M., Paller, K. A., Muthu, K., & Suzuki, S. (2007). Attention induces synchronization-based response gain in steady-state visual evoked potentials. *Nature Neuroscience*, 10(1), 117-125.
- Kipling, R., 1865-1936. (1991). *The jungle book*. New York :Arcade Pub.,
- Koechlin, E., Ody, C., & Kouneiher, F. (2003). The architecture of cognitive control in the human prefrontal cortex. *Science*, 302(5648), 1181-1185.
- Koenig, M. A., Clément, F., & Harris, P. L. (2004). Trust in testimony: Children's use of true and false statements. *Psychological Science*, 15(10), 694-698.
- Kohonen, V., & Service, E. (1995). Is the relation between phonological memory and foreign language learning accounted for by vocabulary acquisition?. *Applied Psycholinguistics*, 16(02), 155-172.
- Korkman, M., Kirk, U., & Kemp, S. (2007). *NEPSY-II*. Madrid: Pearson. [Spanish adaptation by

References

- Universidad de Sevilla, FIVAN and Departamento I_D Pearson Clinical and Talent Assessment (2014). Madrid: Pearson Publ.]
- Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, 33(2), 149.
- Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience*, 5(11), 831-843.
- Kuipers, J.R., Uminski, A., Green, Z., Hughes, D., Aglietti, T. (2017). Word meaning acquisition is reflected in brain potentials of isolated words. *Scientific Reports*, 7, 43341.
- Kutas, M., & Hillyard, S.A., (1980). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science*, 207 (4427), 203–205
- Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, 307(5947), 161-163.
- Kutas, M., & Van Petten, C. (1988). Event-related brain potential studies of language. *Advances in psychophysiology*, 3, 139-187.
- Kutas, M., & Federmeier, K.D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends Cognitive Science*. 4, 463-470.
- 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.
- Laine, M. & Salmelin, R. (2010). Neurocognition of new word learning in the native tongue: lessons from the ancient farming equipment paradigm. *Language Learning*, 60, 25–44.
- Lakatos, P., Shah, A. S., Knuth, K. H., Ulbert, I., Karmos, G., & Schroeder, C. E. (2005). An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *Journal of neurophysiology*, 94(3), 1904-1911.
- Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I., & Schroeder, C. E. (2008). Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science*, 320(5872), 110-113.
- Lakatos, P., Musacchia, G., O'Connell, M. N., Falchier, A. Y., Javitt, D. C., & Schroeder, C. E. (2013). The spectrotemporal filter mechanism of auditory selective attention. *Neuron*, 77(4), 750-761.
- Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics:(de) constructing the N400. *Nature Reviews Neuroscience*, 9(12), 920-933.
- Lee, J. K., Nordahl, C. W., Amaral, D. G., Lee, A., Solomon, M., & Ghetti, S. (2015). Assessing hippocampal development and language in early childhood: Evidence from a new application of the Automatic Segmentation Adapter Tool. *Human Brain Mapping*, 36(11), 4483-4496.
- Li, P., Legault, J., & Litcofsky, K.A. (2014). Neuroplasticity as a function of second language

References

- learning: anatomical changes in the human brain. *Cortex*, 58, 301-324.
- Lieberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological review*, 74(6), 431.
- Linck, J.A., & Weiss, D.J. (2015). Can working memory and inhibitory control predict second language learning in the classroom?. *Sage Open*, 5(4) 2158-2440.
- Lindsay, S., & Gaskell, M. G. (2010). A complementary systems account of word learning in L1 and L2. *Language Learning*, 60, 45-63.
- Lisker, L., & Abramson, A. S. (1971). Distinctive features and laryngeal control. *Language*, 767-785.
- López-Barroso, D., Catani, M., Ripollés, P., Dell'Acqua, F., Rodríguez-Fornells, A., & de Diego-Balaguer, R. (2013). Word learning is mediated by the left arcuate fasciculus. *Proceedings of the National Academy of Sciences*, 110(32), 13168-13173.
- López-Barroso, D., Ripollés, P., Marco-Pallarés, J., Mohammadi, B., Muentel, T. F., Bachoud-Levi, A. C., Rodríguez-Fornells, A., & de Diego-Balaguer, R. (2015). Multiple brain networks underpinning word learning from fluent speech revealed by independent component analysis. *Neuroimage*, 110, 182-193.
- Luck, S. J. (2005). Ten simple rules for designing ERP experiments. *Event-related potentials: A methods handbook*, 262083337.
- Luo, H., & Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron*, 54(6), 1001-1010.
- Luo, H., Liu, Z., & Poeppel, D. (2010). Auditory cortex tracks both auditory and visual stimulus dynamics using low-frequency neuronal phase modulation. *PLoS Biology*, 8(8), e1000445.
- Mack, M.L., Preston, A.R., & Love, B.C. (2020). Ventromedial prefrontal cortex compression during concept learning. *Nat Com*, 11(1), 1-11.
- Malson, L. (1964). *Les Enfants Sauvages: Mythe et réalité*. Paris: Christian Bourgois
- Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, 20(2), 121-157.
- Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive Science*, 14(1), 57-77.
- Markson, L., & Bloom, P. (1997). Evidence against a dedicated system for word learning in children. *Nature*, 385(6619), 813-815.
- Mathewson, K. E., Fabiani, M., Gratton, G., Beck, D. M., & Lleras, A. (2010). Rescuing stimuli from invisibility: Inducing a momentary release from visual masking with pre-target entrainment. *Cognition*, 115(1), 186-191.
- Mattys, S. L., & Jusczyk, P. W. (2001). Phonotactic cues for segmentation of fluent speech by infants. *Cognition*, 78(2), 91-121.

References

- McClelland, J.L., McNaughton, B.L., Randall, C.O. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review* 102, 419–457.
- McLaughlin, J., Osterhout, L., Kim, A. (2004). Neural correlates of second-language word learning: Minimal instruction produces rapid change. *Nature Neuroscience*. 7, 703-704.
- McNealy, K., Mazziotta, J. C., & Dapretto, M. (2006). Cracking the language code: neural mechanisms underlying speech parsing. *Journal of Neuroscience*, 26(29), 7629-7639.
- Medina V., S. M., Paz, R., Roehri, N., Lagarde, S., Pizzo, F., Colombet, B., Bartolomei F., Carron R., & Bénar, C. G. (2018). EpiTools, A software suite for presurgical brain mapping in epilepsy: Intracerebral EEG. *Journal of Neuroscience Methods*, 303, 7-15.
- Mercier, M. R., Bickel, S., Megevand, P., Groppe, D. M., Schroeder, C. E., Mehta, A. D., & Lado, F. A. (2017). Evaluation of cortical local field potential diffusion in stereotactic electroencephalography recordings: A glimpse on white matter signal. *Neuroimage*, 147, 219-232.
- Mestres-Missé, A., Rodriguez-Fornells, A., Münte, T.F. (2007). Watching the brain during meaning acquisition. *Cerebral Cortex*. 17, 1858-1866.
- Mestres-Missé, A., Camara, E., Rodriguez-Fornells, A., Rotte, M., & Münte, T. F. (2008). Functional neuroanatomy of meaning acquisition from context. *Journal of Cognitive Neuroscience*. 20, 2153-2166.
- Mestres-Missé, A., Münte, T. F., & Rodriguez-Fornells, A. (2009). Functional neuroanatomy of contextual acquisition of concrete and abstract words. *Journal of Cognitive Neuroscience*, 21(11), 2154-2171.
- Meyer, T., & Olson, C. R. (2011). Statistical learning of visual transitions in monkey inferotemporal cortex. *Proceedings of the National Academy of Sciences*, 108(48), 19401-19406.
- Midgley, K. J., Holcomb, P. J., & Grainger, J. (2009). Language effects in second language learners and proficient bilinguals investigated with event-related potentials. *Journal of Neurolinguistics*, 22(3), 281-300.
- Mills, D. L., Plunkett, K., Prat, C., & Schafer, G. (2005). Watching the infant brain learn words: Effects of vocabulary size and experience. *Cognitive Development*, 20(1), 19-31.
- Milner, A. D., & Goodale, M. A. (2008). Two visual systems re-viewed. *Neuropsychologia*, 46(3), 774-785.
- Moore, C., Angelopoulos, M., & Bennett, P. (1999). Word learning in the context of referential and salience cues. *Developmental Psychology*, 35(1), 60.
- Morgan, S. T., Hansen, J. C., & Hillyard, S. A. (1996). Selective attention to stimulus location modulates the steady-state visual evoked potential. *Proceedings of the National Academy of Sciences*, 93(10), 4770-4774.
- Morgan-Short, K. (2014). Electrophysiological approaches to understanding second language acquisition: A field reaching its potential. *Annual Review of Applied Linguist*, 34, 15-36.

References

- Mumford, D. (1992). On the computational architecture of the neocortex. *Biological cybernetics*, 66(3), 241-251.
- Murray, L. J., & Ranganath, C. (2007). The dorsolateral prefrontal cortex contributes to successful relational memory encoding. *Journal of Neuroscience*, 27(20), 5515-5522.
- Nathan, M. J., & Del Pinal, G. (2017). The future of cognitive neuroscience? Reverse inference in focus. *Philosophy Compass*, 12(7), e12427.
- Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development*, 1-135.
- Nettl, B. (2000). An ethnomusicologist contemplates universals in musical sound and musical culture. *The Origins of Music*, 3(2), 463-472.
- Newman, R., Ratner, N. B., Jusczyk, A. M., Jusczyk, P. W., & Dow, K. A. (2006). Infants' early ability to segment the conversational speech signal predicts later language development: A retrospective analysis. *Developmental Psychology*, 42(4), 643.
- Newport, E. L. (1990). Maturation constraints on language learning. *Cognitive Science*, 14(1), 11-28.
- Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, 48(2), 127-162.
- Ngon, C., Martin, A., Dupoux, E., Cabrol, D., Dutat, M., & Peperkamp, S. (2013). (Non) words,(non) words,(non) words: evidence for a protolexicon during the first year of life. *Developmental Science*, 16(1), 24-34.
- Norman-Haignere, S., Kanwisher, N. G., & McDermott, J. H. (2015). Distinct cortical pathways for music and speech revealed by hypothesis-free voxel decomposition. *Neuron*, 88(6), 1281-1296.
- Norman-Haignere, S., Feather, J., Brunner, P., Ritaccio, A., McDermott, J. H., Schalk, G., & Kanwisher, N. (2019). Intracranial recordings from human auditory cortex reveal a neural population selective for musical song. *BioRxiv*, 696161.
- Nozaradan, S., Peretz, I., Missal, M., & Mouraux, A. (2011). Tagging the neuronal entrainment to beat and meter. *Journal of Neuroscience*, 31(28), 10234-10240.
- Nozaradan, S., Peretz, I., & Mouraux, A. (2012). Selective neuronal entrainment to the beat and meter embedded in a musical rhythm. *Journal of Neuroscience*, 32(49), 17572-17581.
- Nozaradan, S. (2014). Exploring how musical rhythm entrains brain activity with electroencephalogram frequency-tagging. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1658), 20130393.
- Nozaradan, S., Mouraux, A., & Cousineau, M. (2017). Frequency tagging to track the neural processing of contrast in fast, continuous sound sequences. *Journal of Neurophysiology*, 118(1), 243-253.
- Obleser, J., & Kayser, C. (2019). Neural entrainment and attentional selection in the listening

References

- brain. *Trends in Cognitive Sciences*, 23(11), 913-926.
- Oever, S., Schroeder, C. E., Poeppel, D., Van Atteveldt, N., Mehta, A. D., Mégevand, P., Zion-Golumbi E., & Zion-Golumbic, E. (2017). Low-frequency cortical oscillations entrain to subthreshold rhythmic auditory stimuli. *Journal of Neuroscience*, 37(19), 4903-4912.
- Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*.
- Ordin, M., Polyanskaya, L., Soto, D., & Molinaro, N. (2020). Electrophysiology of statistical learning: Exploring the online learning process and off-line learning product. *European Journal of Neuroscience*, 51(9), 2008-2022.
- Ortiz, E., Stingl, K., Münßinger, J., Braun, C., Preissl, H., & Belardinelli, P. (2012). Weighted phase lag index and graph analysis: preliminary investigation of functional connectivity during resting state in children. *Computational and Mathematical Methods in Medicine*, 2012.
- Overath, T., McDermott, J. H., Zarate, J. M., & Poeppel, D. (2015). The cortical analysis of speech-specific temporal structure revealed by responses to sound quilts. *Nature Neuroscience*, 18(6), 903.
- Palermo, R., & Rhodes, G. (2007). Are you always on my mind? A review of how face perception and attention interact. *Neuropsychologia*, 45(1), 75-92.
- Paller, K. A., Voss, J. L., & Boehm, S. G. (2007). Validating neural correlates of familiarity. *Trends in Cognitive Sciences*, 11(6), 243-250.
- Panda, E. J., Emami, Z., Valiante, T. A., & Pang, E. W. (2020). EEG Phase Synchronization during Semantic Unification Relates to Individual Differences in Children's Vocabulary Skill. *Developmental Science*, e12984.
- Papagno, C., Miracapillo, C., Casarotti, A., Romero Lauro, L. J., Castellano, A., Falini, A., Casaceli, G., Fava, E., & Bello, L. (2011). What is the role of the uncinate fasciculus? Surgical removal and proper name retrieval. *Brain*, 134(2), 405-414.
- Parasuraman, R. (1980). Effects of information processing demands on slow negative shift latencies and N100 amplitude in selective and divided attention. *Biological Psychology*, 11(3-4), 217-233.
- Pasley, B. N., David, S. V., Mesgarani, N., Flinker, A., Shamma, S. A., Crone, N. E., Knight, R. T. & Chang, E. F. (2012). Reconstructing speech from human auditory cortex. *PLoS Biology*, 10(1), e1001251.
- Pasquini, E. S., Corriveau, K. H., Koenig, M., & Harris, P. L. (2007). Preschoolers monitor the relative accuracy of informants. *Developmental Psychology*, 43(5), 1216.
- Paulesu, E., Vallar, G., Berlingeri, M., Signorini, M., Vitali, P., Burani, C., & Fazio, F. (2009). Supercalifragilisticexpialidocious: how the brain learns words never heard before. *Neuroimage*. 45(4), 1368-1377.
- Peelle, J. E., Gross, J., & Davis, M. H. (2013). Phase-locked responses to speech in human auditory

References

- cortex are enhanced during comprehension. *Cerebral Cortex*, 23(6), 1378-1387.
- Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009a). Learning in reverse: Eight-month-old infants track backward transitional probabilities. *Cognition*, 113(2), 244-247.
- Pelucchi, B., Hay, J. F., & Saffran, J. R. (2009b). Statistical learning in a natural language by 8-month-old infants. *Child Development*, 80(3), 674-685.
- Penny, W., Friston, K., Ashburner, J., Kiebel, S., & Nichols, T., 2006. *Statistical parametric mapping. Annals Functional Brain Images*, London, UK.
- Penaloza, C., Mirman, D., Tuomiranta, L., Benetello, A., Heikius, I. M., Järvinen, S., Majos, M.C.; Juncadella, M., Laine, M., Martin, N., & Rodríguez-Fornells, A. (2016). Novel word acquisition in aphasia: Facing the word-referent ambiguity of natural language learning contexts. *Cortex*, 79, 14-31.
- Peñaloza, C., Mirman, D., Cardona, P., Juncadella, M., Martin, N., Laine, M., & Rodríguez-Fornells, A. (2017). Cross-situational word learning in aphasia. *Cortex*, 93, 12-27.
- Perfetti, C.A., Wlotko, E.W., & Hart, L.A. (2005). Word Learning and Individual differences in word learning reflected in event-related potentials. *Journal of Experimental Psychology in Learning, Memory & Cognition*. 31, 1281–1292.
- Perfetti, C. (2007) Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading* 11, 357-383.
- Perruchet, P., & Vinter, A. (1998). PARSER: A model for word segmentation. *Journal of Memory and Language*, 39(2), 246-263.
- Petersen, S.E., Fox, P., Postner, M., Mintun, M., Raichle, M.E. (1988). Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature* 331, 585–589.
- Pettersen, C. V., Kutas, M., Kluender, R., Mitchiner, M., & McIsaac, H. (1991). Fractionating the word repetition effect with event-related potentials. *Journal of Cognitive Neuroscience*, 3(2), 131-150.
- Piaget, J. (1954). *The construction of reality in the child*. New York: Free Press
- Piai, V., Anderson, K. L., Lin, J. J., Dewar, C., Parvizi, J., Dronkers, N. F., & Knight, R. T. (2016). Direct brain recordings reveal hippocampal rhythm underpinnings of language processing. *Proceedings of the National Academy of Sciences*, 113(40), 11366-11371.
- Picton, T. W., John, M. S., Dimitrijevic, A., & Purcell, D. (2003). Human auditory steady-state responses: Respuestas auditivas de estado estable en humanos. *International Journal of Audiology*, 42(4), 177-219.
- Pillai, J. J., Araque, J. M., Allison, J. D., Sethuraman, S., Loring, D. W., Thiruvaiyaru, D., & Lavin, T. (2003). Functional MRI study of semantic and phonological language processing in bilingual subjects: preliminary findings. *NeuroImage*, 19(3), 565-576.
- Pliatsikas, C., Johnstone, T., & Marinis, T. (2014). Grey matter volume in the cerebellum is related

References

- to the processing of grammatical rules in a second language: a structural voxel-based morphometry study. *The Cerebellum*, 13(1), 55-63.
- Plunkett, K. (1993). Lexical segmentation and vocabulary growth in early language acquisition. *Journal of Child Language*, 20(1), 43-60.
- Plunkett, K. (1997). Theories of early language acquisition. *Trends in Cognitive Sciences*, 1(4), 146-153.
- Poeppl, D., & Hickok, G. (2004). Towards a new functional anatomy of language. *Cognition*, 92(1-2): 1-12.
- Poldrack, R. A. (2000). Imaging brain plasticity: conceptual and methodological issues—a theoretical review. *Neuroimage*, 12(1), 1-13.
- Poldrack, R. A., & Packard, M. G. (2003). Competition among multiple memory systems: converging evidence from animal and human brain studies. *Neuropsychologia*, 41(3), 245-251.
- Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage*, 59(3), 2142-2154.
- Price, C. J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *Neuroimage*, 62(2), 816-847.
- Qi, Z., Beach, S.D., Finn, A.S., Minas, J., Goetz, C., Chan, B., & Gabrieli, J.D. (2017). Native-language N400 and P600 predict dissociable language-learning abilities in adults. *Neuropsychologia*, 98, 177-191.
- Quine, W.V.O., 1960. *Word and Object*. MIT Press, Cambridge, MA.
- Radvansky, G.A. & Zacks, J.M. (2017). Event boundaries in memory and cognition. *Current Opinion in Behavioral Sciences*, 17, 133-140.
- Raichle, M. E., Fiez, J. A., Videen, T. O., MacLeod, A. M. K., Pardo, J. V., Fox, P. T., & Petersen, S. E. (1994). Practice-related changes in human brain functional anatomy during nonmotor learning. *Cerebral Cortex*, 4(1), 8-26.
- Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences*, 98(2), 676-682.
- Rasamimanana, M., Barbaroux, M., Colé, P., & Besson, M. (2020). Semantic compensation and novel word learning in university students with dyslexia. *Neuropsychologia*, 139, 107358.
- Räsänen, O., & Rasilo, H. (2015). A joint model of word segmentation and meaning acquisition through cross-situational learning. *Psychological Review*, 122(4), 792.
- Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nature Neuroscience*, 12(6), 718-724.

References

- Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, 6(6), 855-863.
- Reber, A.S. & Allen, R. (2000). Individual Differences in Implicit Learning Implications for the Evolution of Consciousness. *Individual Differences in Conscious Experience*, 20, 227.
- Regan, D. (1982). Comparison of transient and steady-state methods. *Annals of the New York Academy of Sciences*, 388(1): 45-71.
- Reiterer, S. M., Pereda, E., & Bhattacharya, J. (2011). On a possible relationship between linguistic expertise and EEG gamma band phase synchrony. *Frontiers in Psychology*, 2, 334.
- Ríos-López, P., Molinaro, N., Bourguignon, M., & Lallier, M. (2020). Development of neural oscillatory activity in response to speech in children from 4 to 6 years old. *Developmental Science*, e12947.
- Ripollés, P., Marco-Pallarés, J., Hielscher, U., Mestres-Missé, A., Tempelmann, C., Heinze, H. J., & Noesselt, T. (2014). The role of reward in word learning and its implications for language acquisition. *Current Biology*. 24(21), 2606-2611.
- Ripollés, P., Marco-Pallares, J., Alicart, H., Tempelmann, C., Rodriguez-Fornells, A., & Noesselt, T. (2016). Intrinsic monitoring of learning success facilitates memory encoding via the activation of the SN/VTA-Hippocampal loop. *Elife*, 5, e17441.
- Rizzuto, R., & Pozzan, T. (2006). Microdomains of intracellular Ca²⁺: molecular determinants and functional consequences. *Physiological Reviews*, 86(1), 369-408.
- Rodríguez-Fornells, A., Cunillera, T., Mestres-Missé, A., & de Diego-Balaguer, R. (2009). Neurophysiological mechanisms involved in language learning in adults. *Philosophical Transactions of the Royal Society: Biological Sciences*, 364(1536), 3711-3735.
- Rodríguez-Fornells, A., Kramer, U., Lorenzo-Seva, U., Festman, J., & Münte, T. F. (2012). Self-assessment of individual differences in language switching. *Frontiers in Psychology*, 2, 388.
- Roelofs, A., & Piai, V. (2011). Attention demands of spoken word planning: A review. *Frontiers in Psychology*, 2, 307.
- Rugg, M. D., Doyle, M. C., & Wells, T. (1995). Word and nonword repetition within-and across-modality: An event-related potential study. *Journal of Cognitive Neuroscience*, 7(2), 209-227.
- Rugg, M. D., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in Cognitive Sciences*, 11(6), 251-257.
- Ryan, J. D., Althoff, R. R., Whitlow, S., & Cohen, N. J. (2000). Amnesia is a deficit in relational memory. *Psychological Science*, 11(6), 454-461.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274(5294), 1926-1928.
- Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science*, 8(2),

References

- 101-105.
- Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, 70(1), 27-52.
- Saffran, J. R. (2003). Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science*, 12(4), 110-114.
- Sagi, Y., Tavor, I., Hofstetter, S., Tzur-Moryosef, S., Blumenfeld-Katzir, T., & Assaf, Y. (2012). Learning in the fast lane: new insights into neuroplasticity. *Neuron*, 73(6), 1195-1203.
- Sahni, S. D., Seidenberg, M. S., & Saffran, J. R. (2010). Connecting cues: Overlapping regularities support cue discovery in infancy. *Child Development*, 81(3), 727-736.
- Saleh, M., Reimer, J., Penn, R., Ojakangas, C. L., & Hatsopoulos, N. G. (2010). Fast and slow oscillations in human primary motor cortex predict oncoming behaviorally relevant cues. *Neuron*, 65(4), 461-471.
- Sanders, L. D., & Neville, H. J. (2000). Lexical, syntactic, and stress-pattern cues for speech segmentation. *Journal of Speech, Language, and Hearing Research*, 43(6), 1301-1321.
- Sanders, L. D., Newport, E. L., & Neville, H. J. (2002). Segmenting nonsense: an event-related potential index of perceived onsets in continuous speech. *Nature Neuroscience*, 5(7), 700-703.
- Sanders, L. D., & Neville, H. J. (2003). An ERP study of continuous speech processing: I. Segmentation, semantics, and syntax in native speakers. *Cognitive Brain Research*, 15(3), 228-240.
- Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M. S., Umarova, R., Musso, M., Glauche, V., Abel, S., Huber, W., Rijntjes, M., Hennig, J., & Weiller, C. (2008). Ventral and dorsal pathways for language. *Proceedings of the National Academy of Sciences*, 105(46), 18035-18040.
- Sauseng, P., Klimesch, W., Schabus, M., & Doppelmayr, M. (2005). Fronto-parietal EEG coherence in theta and upper alpha reflect central executive functions of working memory. *International Journal of Psychophysiology*, 57(2), 97-103.
- Schapiro, A. C., Kustner, L. V., & Turk-Browne, N. B. (2012). Shaping of object representations in the human medial temporal lobe based on temporal regularities. *Current biology*, 22(17), 1622-1627.
- Schapiro, A. C., Gregory, E., Landau, B., McCloskey, M., & Turk-Browne, N. B. (2014). The necessity of the medial temporal lobe for statistical learning. *Journal of Cognitive Neuroscience*, 26(8), 1736-1747.
- Schapiro, A., & Turk-Browne, N. (2015). Statistical learning. *Brain Mapping*, 3, 501-506.
- Schapiro, A. C., Turk-Browne, N. B., Botvinick, M. M., & Norman, K. A. (2017). Complementary learning systems within the hippocampus: a neural network modelling approach to reconciling episodic memory with statistical learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1711), 20160049.

References

- Schlichting, M. L., Guarino, K. F., Schapiro, A. C., Turk-Browne, N. B., & Preston, A. R. (2017). Hippocampal structure predicts statistical learning and associative inference abilities during development. *Journal of Cognitive Neuroscience*, 29(1), 37-51.
- Schmahmann, J. D., Smith, E. E., Eichler, F. S., & Filley, C. M. (2008). Cerebral white matter: neuroanatomy, clinical neurology, and neurobehavioral correlates. *Annals of the New York Academy of Sciences*, 1142, 266.
- Schön, D., Boyer, M., Moreno, S., Besson, M., Peretz, I., & Kolinsky, R. (2008). Songs as an aid for language acquisition. *Cognition*, 106(2), 975-983.
- Schroeder, C. E., & Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends in Neurosciences*, 32(1), 9-18.
- Scott-Van Zeeland, A. A., McNealy, K., Wang, A. T., Sigman, M., Bookheimer, S. Y., & Dapretto, M. (2010). No neural evidence of statistical learning during exposure to artificial languages in children with autism spectrum disorders. *Biological psychiatry*, 68(4), 345-351.
- Service, E. (1992). Phonology, working memory, and foreign-language learning. *The Quarterly Journal of Experimental Psychology Section A*, 45(1), 21-50.
- Sestieri, C., Shulman, G.L., Corbetta, M. (2017). The contribution of the human posterior parietal cortex to episodic memory. *Nature Review of Neuroscience*. 18, 183.
- Shao, Z., Janse, E., Visser, K., & Meyer, A.S. (2014). What do verbal fluency tasks measure? Predictors of verbal fluency performance in older adults. *Frontiers in Psychology*, 5, 772.
- Sharon, T., Moscovitch, M., & Gilboa, A. (2011) Rapid neocortical acquisition of long-term arbitrary associations independent of the hippocampus. *Proceedings of the National Academy of Sciences*. 108.3. 1146-1151.
- Shipley, E. F., & Shepperson, B. (1990). Countable entities: Developmental changes. *Cognition*, 34(2), 109-136.
- Shukla, M., White, K. S., & Aslin, R. N. (2011). Prosody guides the rapid mapping of auditory word forms onto visual objects in 6-mo-old infants. *Proceedings of the National Academy of Sciences*, 108(15), 6038-6043.
- Siegelman, N. (2020). Statistical learning abilities and their relation to language. *Language and Linguistics Compass*, 14(3), e12365.
- Siegelman, N., Bogaerts, L., Christiansen, M. H., & Frost, R. (2017). Towards a theory of individual differences in statistical learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1711), 20160059.
- Singh, L., Steven Reznick, J., & Xuehua, L. (2012). Infant word segmentation and childhood vocabulary development: a longitudinal analysis. *Developmental Science*, 15(4), 482-495.
- Sliwinska, M.W., Violante, I.R., Wise, R.J., Leech, R., Devlin, J.T., Geranmayeh, F., & Hampshire, A. (2017). Stimulating multiple-demand cortex enhances vocabulary learning. *Journal of Neuroscience*. 37(32), 7606-7618.

References

- Smith, L. B. (2000). From knowledge to knowing: Real progress in the study of infant categorization. *Infancy*, 1(1), 91-97.
- Smith, Z. M., Delgutte, B., & Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, 416(6876), 87-90.
- Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, 106(3), 1558-1568.
- Smith, L. B., & Jones, S. S. (2011). Symbolic play connects to language through visual object recognition. *Developmental Science*, 14(5), 1142-1149.
- Smith, L. B., Suanda, S. H., & Yu, C. (2014). The unrealized promise of infant statistical word-referent learning. *Trends in Cognitive Sciences*, 18(5), 251-258.
- Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6(2), 174.
- Snow, C. E., & Hoefnagel-Höhle, M. (1978). The critical period for language acquisition: Evidence from second language learning. *Child Development*, 1114-1128.
- Sohoglu, E., Peelle, J. E., Carlyon, R. P., & Davis, M. H. (2012). Predictive top-down integration of prior knowledge during speech perception. *Journal of Neuroscience*, 32(25), 8443-8453.
- Sonty, S. P., Mesulam, M. M., Weintraub, S., Johnson, N. A., Parrish, T. B., & Gitelman, D. R. (2007). Altered effective connectivity within the language network in primary progressive aphasia. *Journal of Neuroscience*, 27(6), 1334-1345.
- Soskey, L., Holcomb, P.J., & Midgley, K.J. (2016). Language effects in second-language learners: A longitudinal electrophysiological study of Spanish classroom learning. *Brain Research*. 1646, 45–52.
- Southwell, R., Baumann, A., Gal, C., Barascud, N., Friston, K., & Chait, M. (2017). Is predictability salient? A study of attentional capture by auditory patterns. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1714), 20160105.
- Speer, N. K., Zacks, J. M., & Reynolds, J. R. (2007). Human brain activity time-locked to narrative event boundaries. *Psychological Science*, 18(5), 449-455.
- Squire, L. R., & Knowlton, B. J. (2000). The medial temporal lobe, the hippocampus, and the memory systems of the brain. *The New Cognitive Neurosciences*, 2, 756-776.
- Squire, L. R. (2004). Memory systems of the brain: a brief history and current perspective. *Neurobiology of Learning and Memory*, 82(3), 171-177.
- Stam, C. J., Nolte, G., & Daffertshofer, A. (2007). Phase lag index: assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources. *Human Brain Mapping*, 28(11), 1178-1193.
- Staresina, B. P., & Davachi, L. (2009). Mind the gap: binding experiences across space and time in the human hippocampus. *Neuron*, 63(2), 267-276.

References

- Stein, M., Dierks, T., Brandeis, D., Wirth, M., Strik, W., & König, T. (2006). Plasticity in the adult language system: A longitudinal electrophysiological study on Second Language Learning. *NeuroImage*, 33, 774–783.
- Steriade, M. (2001). Impact of network activities on neuronal properties in corticothalamic systems. *Journal of Neurophysiology*, 86(1), 1-39.
- Stevens, K. N. (2002). Toward a model for lexical access based on acoustic landmarks and distinctive features. *The Journal of the Acoustical Society of America*, 111(4), 1872-1891.
- Strijkers, K., Holcomb, P. J., & Costa, A. (2011). Conscious intention to speak proactively facilitates lexical access during overt object naming. *Journal of Memory and Language*, 65(4), 345-362.
- Strijkers, K., Bertrand, D., & Grainger, J. (2015). Seeing the same words differently: The time course of automaticity and top-down intention in reading. *Journal of Cognitive Neuroscience*, 27(8), 1542-1551.
- Studdert-Kennedy, M. (1987). The phoneme as a perceptuomotor structure. Haskins Laboratories: *Status Report on Speech Research*, SR, 91, 45-57.
- Stuellein, N., Radach, R. R., Jacobs, A. M., & Hofmann, M. J. (2016). No one way ticket from orthography to semantics in recognition memory: N400 and P200 effects of associations. *Brain Research*, 1639, 88-98.
- Stuss, D. T., Alexander, M. P., Shallice, T., Picton, T. W., Binns, M. A., Macdonald, R., Borowiec, A., & Katz, D. I. (2005). Multiple frontal systems controlling response speed. *Neuropsychologia*, 43(3), 396-417.
- Stuss, D. T. (2011). Functions of the frontal lobes: relation to executive functions. *Journal of the International Neuropsychological Society: JINS*, 17(5), 759.
- Swingle, D. (2005). Statistical clustering and the contents of the infant vocabulary. *Cognitive Psychology*, 50(1), 86-132.
- Tagarelli, K.M., Shattuck, K.F., Turkeltaub, P.E., & Ullman, M.T. (2019). Language learning in the adult brain: A neuroanatomical meta-analysis of lexical and grammatical learning. *NeuroImage*, 193, 178-200.
- Tager-Flusberg, H., Boshart, J., & Baron-Cohen, S. (1998). Reading the windows to the soul: Evidence of domain-specific sparing in Williams syndrome. *Journal of Cognitive Neuroscience*, 10(5), 631-639.
- Talsma, D., Senkowski, D., Soto-Faraco, S., & Woldorff, M. G. (2010). The multifaceted interplay between attention and multi-sensory integration. *Trends in Cognitive Sciences*, 14(9), 400-410.
- Tamminen, J., & Gaskell, M.G. (2013). Novel word integration in the mental lexicon: Evidence from unmasked and masked semantic priming. *Quarterly Journal of Experimental Psychology*, 66, 1001-1025.
- Taylor, B. K., Embury, C. M., Heinrichs-Graham, E., Frenzel, M. R., Eastman, J. A., Wiesman, A. I.,

References

- Wang, Y.P., Calhoun, V. D., Stephen, J. M., & Wilson, T. W. (2020). Neural oscillatory dynamics serving abstract reasoning reveal robust sex differences in typically-developing children and adolescents. *Developmental Cognitive Neuroscience*, 42, 100770.
- Teinonen, T., Fellman, V., Näätänen, R., Alku, P., & Huotilainen, M. (2009). Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neuroscience*, 10(1), 1-8.
- Tham, E. K., Lindsay, S., & Gaskell, M. G. (2015). Markers of automaticity in sleep-associated consolidation of novel words. *Neuropsychologia*, 71, 146-157.
- Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: use of stress and statistical cues to word boundaries by 7-to 9-month-old infants. *Developmental Psychology*, 39(4), 706.
- Thiessen, E. D. (2010). Effects of visual information on adults' and infants' auditory statistical learning. *Cognitive Science*, 34(6), 1093-1106.
- Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: a reevaluation. *Proceedings of the National Academy of Sciences*, 94(26), 14792-14797.
- Thut, G., Schyns, P., & Gross, J. (2011). Entrainment of perceptually relevant brain oscillations by non-invasive rhythmic stimulation of the human brain. *Frontiers in Psychology*, 2, 170.
- Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child Development*, 1454-1463.
- Tomasello, M., & Barton, M. E. (1994). Learning words in nonostensive contexts. *Developmental Psychology*, 30(5), 639.
- Tomasello, M. (1995). Joint attention as social cognition. Joint attention: Its origins and role in development, 103130.
- Tomasello, M. (2000). First steps toward a usage-based theory of language acquisition. *Cognitive Linguistics*, 11(1/2), 61-82.
- Torkildsen, J.V.K., Svangstu, J.M., Hansen, H.F., Smith, L., Simonsen, H.G., Moen, I., & Lindgren, M. (2008). Productive vocabulary size predicts event-related potential correlates of fast mapping in 20-Month-Olds. *Journal of Cognitive Neuroscience*. 20, 1266–1282.
- Toro-Soto, J. M., & Trobalón, J. B. (2005). Statistical computations over a speech stream in a rodent. *Perception & Psychophysics*, 67(5), 867-875.
- Toro-Soto, J. M., Rodríguez-Fornells, A., & Sebastián-Gallés, N. (2007). Stress placement and word segmentation by Spanish speakers. *Psicológica*, 28(2), 167-176.
- Tort, A. B., Komorowski, R. W., Manns, J. R., Kopell, N. J., & Eichenbaum, H. (2009). Theta-gamma coupling increases during the learning of item-context associations. *Proceedings of the National Academy of Sciences*, 106(49), 20942-20947.
- Tremblay, P., & Dick, A. S. (2016). Broca and Wernicke are dead, or moving past the classic model of language neurobiology. *Brain and Language*, 162, 60-71.

References

- Tremblay, P., Baroni, M., & Hasson, U. (2013). Processing of speech and non-speech sounds in the supratemporal plane: auditory input preference does not predict sensitivity to statistical structure. *Neuroimage*, 66, 318-332.
- Tse, D., Langston, R.F., Kakeyama, M., Bethus, I., Spooner, P.A., Wood, E.R., Witter, M.P., & Morris, R.G.M. (2007). Schemas and Memory Consolidation. *Science*, 316, 76–82
- Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, 134(4), 552.
- Turk-Browne, N. B., Scholl, B. J., Chun, M. M., & Johnson, M. K. (2009). Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *Journal of Cognitive Neuroscience*, 21(10), 1934-1945.
- Turk-Browne, N. B., Scholl, B. J., Johnson, M. K., & Chun, M. M. (2010). Implicit perceptual anticipation triggered by statistical learning. *Journal of Neuroscience*, 30(33), 11177-11187.
- Turken, A. U., & Dronkers, N. F. (2011). The neural architecture of the language comprehension network: converging evidence from lesion and connectivity analyses. *Frontiers in Systems Neuroscience*. 2011; 5: 1.
- Tzourio, N., El Massioui, F., Crivello, F., Joliot, M., Renault, B., & Mazoyer, B. (1997). Functional anatomy of human auditory attention studied with PET. *Neuroimage*, 5(1), 63-77.
- Ullman, M.T. (2001). A Neurocognitive perspective on language: The declarative/procedural model. *Nature Review of Neuroscience*. 2, 717–726.
- Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, 92(1-2), 231-270.
- Van Diepen, R. M., Foxe, J. J., & Mazaheri, A. (2019). The functional role of alpha-band activity in attentional processing: the current zeitgeist and future outlook. *Current Opinion in Psychology*, 29, 229-238.
- van Kesteren, M.T., Ruiters, D.J., Fernández, G., & Henson, R.N. (2012). How schema and novelty augment memory formation. *Trends in Neuroscience*. 35, 211-219.
- Vander Ghinst, M., Bourguignon, M., Niesen, M., Wens, V., Hassid, S., Choufani, G., Joumäki, V., Hari, R., Goldman, S., & De Tiège, X. (2019). Cortical tracking of speech-in-noise develops from childhood to adulthood. *Journal of Neuroscience*, 39(15), 2938-2950.
- Vincent, J. L., Kahn, I., Snyder, A. Z., Raichle, M. E., & Buckner, R. L. (2008). Evidence for a frontoparietal control system revealed by intrinsic functional connectivity. *Journal of Neurophysiology*, 100(6), 3328-3342.
- Vinck, M., Oostenveld, R., Van Wingerden, M., Battaglia, F., & Pennartz, C. M. (2011). An improved index of phase-synchronization for electrophysiological data in the presence of volume-conduction, noise and sample-size bias. *Neuroimage*, 55(4), 1548-1565.
- Voss, J. L., & Federmeier, K. D. (2011). FN400 potentials are functionally identical to N400 potentials and reflect semantic processing during recognition testing. *Psychophysiology*,

References

- 48(4), 532-546.
- Voytek, B., Canolty, R. T., Shestyuk, A., Crone, N., Parvizi, J., & Knight, R. T. (2010). Shifts in gamma phase–amplitude coupling frequency from theta to alpha over posterior cortex during visual tasks. *Frontiers in Human Neuroscience*, 4, 191.
- Vygotsky, L.S. (1980). *Mind in Society: The development of Higher Psychological Processes*. London, England. Harvard University Press.
- Wang, J., Gao, D., Li, D., Desroches, A. S., Liu, L., & Li, X. (2014). Theta–gamma coupling reflects the interaction of bottom-up and top-down processes in speech perception in children. *Neuroimage*, 102, 637-645.
- Warren, D. E., Roembke, T. C., Covington, N. V., McMurray, B., & Duff, M. C. (2019). Cross-situational statistical learning of new words despite bilateral hippocampal damage and severe amnesia. *Frontiers in Human Neuroscience*, 13.
- Warren, D. H., Welch, R. B., & McCarthy, T. J. (1981). The role of visual-auditory “compellingness” in the ventriloquism effect: Implications for transitivity among the spatial senses. *Perception & Psychophysics*, 30(6), 557-564.
- Weber-Fox, C., & Neville, H. J. (2001). Sensitive periods differentiate processing of open-and closed-class words. *Journal of Speech, Language, and Hearing Research*, 44(6):1338-53.
- Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science*, 24(11), 2143-2152.
- Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7(1), 49-63.
- Wernicke, C. (1874). *Der aphasische Symptomencomplex: eine psychologische Studie auf anatomischer Basis*. Cohn & Weigert.
- Wilson, M., & Emmorey, K. (2006). Comparing sign language and speech reveals a universal limit on short-term memory capacity. *Psychological Science -cambridge-*, 17(8), 682.
- Wilson, S. M., Saygin, A. P., Sereno, M. I., & Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nature Neuroscience*, 7(7), 701-702.
- Wise, R. J., Scott, S. K., Blank, S. C., Mummery, C. J., Murphy, K., & Warburton, E. A. (2001). Separate neural subsystems within Wernicke's area'. *Brain*, 124(1), 83-95.
- Woldorff, M. G., Gallen, C. C., Hampson, S. A., Hillyard, S. A., Pantev, C., Sobel, D., & Bloom, F. E. (1993). Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proceedings of the National Academy of Sciences*, 90(18), 8722-8726.
- Worden, M. S., Foxe, J. J., Wang, N., & Simpson, G. V. (2000). Anticipatory biasing of visuospatial attention indexed by retinotopically specific α -band electroencephalography increases over occipital cortex. *Journal of Neuroscience*, 20(6), RC63-RC63.
- Wöstmann, M., Fiedler, L., & Obleser, J. (2017). Tracking the signal, cracking the code: Speech and speech comprehension in non-invasive human electrophysiology. *Language, Cognition*

References

- and Neuroscience*, 32(7), 855-869.
- Xia, M., Wang, J., & He, Y. (2013). BrainNet Viewer: a network visualization tool for human brain connectomics. *PLoS One*, 8(7), e68910.
- Yang, J., Gates, K. M., Molenaar, P., & Li, P. (2015). Neural changes underlying successful second language word learning: An fMRI study. *Journal of Neurolinguistics*, 33, 29-49.
- Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science*, 18(5), 414-420.
- Yu, C., & Smith, L. B. (2011). What you learn is what you see: using eye movements to study infant cross-situational word learning. *Developmental Science*, 14(2), 165-180.
- Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition*, 125(2), 244-262.
- Yum, Y.N., Midgley, J.K., Holcomb, P.J., Grainger, J. (2014). An ERP study on initial second language vocabulary learning. *Psychophysiology* 51, 364–373.
- Yurovsky, D., Yu, C., & Smith, L. B. (2012). Statistical speech segmentation and word learning in parallel: scaffolding from child-directed speech. *Frontiers in Psychology*, 3, 374.
- Zaccarella, E., & Friederici, A. D. (2017). The neurobiological nature of syntactic hierarchies. *Neuroscience & Biobehavioral Reviews*, 81, 205-212.
- Zacks, J. M., Braver, T. S., Sheridan, M. A., Donaldson, D. I., Snyder, A. Z., Ollinger, J. M., Buckner R.L, & Raichle, M. E. (2001). Human brain activity time-locked to perceptual event boundaries. *Nature Neuroscience*, 4(6), 651-655.
- Zacks, J. M., & Swallow, K. M. (2007). Event segmentation. *Current Directions in Psychological Science*, 16(2), 80-84.
- Zeng, F. G., Nie, K., Stickney, G. S., Kong, Y. Y., Vongphoe, M., Bhargave, A., Wei, C., & Cao, K. (2005). Speech recognition with amplitude and frequency modulations. *Proceedings of the National Academy of Sciences*, 102(7), 2293-2298.
- Zoefel, B., Ten Oever, S., & Sack, A. T. (2018). The involvement of endogenous neural oscillations in the processing of rhythmic input: More than a regular repetition of evoked neural responses. *Frontiers in Neuroscience*, 12, 95.
- Züst, M. A., Ruch, S., Wiest, R., & Henke, K. (2019). Implicit vocabulary learning during sleep is bound to slow-wave peaks. *Current Biology*, 29(4), 541-553.