The Influence of Literacy on Speech

The Orthographic Consistency Effect in Auditory Language Perception and Language Production

Alberto Furgoni, M.A.

Supervised by

Prof. Clara D. Martin & Dr. Antje Stoehr

eman ta zabal zazu

Universidad del País Vasco
Euskal Herriko Unibertsitatea

2022

eman ta zabal zazu

Universidad
del País Vasco

Euskal Herriko
Unibertsitatea

bcbl

BASQUE CENTER
ON COGNITION, BRAIN
AND LANGUAGE

The Influence of Literacy on Speech

The Orthographic Consistency Effect in auditory Language Perception and Language Production

Doctoral dissertation by

Alberto Furgoni, M.A.

Thesis submitted for the degree of

Doctor of Philosophy

Supervised by

Prof. Clara D. Martin & Dr. Antje Stoehr

Basque Center on Cognition, Brain and Language (BCBL)

Euskal Herriko Unibertsitatea – Universidad del País Vasco (EHU/UPV)

Donostia-San Sebastián, 2021

The project that gave rise to these results received the support of a fellowship from ”la Caixa” Foundation (ID 100010434). The fellowship code is LCF/BQ/IN18/11660068. This project has also received funding from the European Union´s Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 713673.[1]

The research was supported by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation program (grant agreement No. 819093). [2]

Alberto Furgoni

BCBL – Basque Center on Cognition, Brain and Language
Paseo Mikeletegi, 69,
20009 - Donsotia-San Sebastián, Spain

2022

---

[1] Awarded to Alberto Furgoni

[2] Awarded to Clara D. Martin

A mia madre

To my mother

Acknowledgements (professional)

I think that doing a PhD is not only a steep career path but also an emotional path. It was, at least in my case, also an emotional path since I experienced all kinds of emotions in these last three years and a half: some of them were directly related to the PhD itself, others were not. Nevertheless, I will remember these "PhD years" especially thanks to the many people who were part of this journey, both professionally and personally.

First, I want to thank my supervisors Dr. Clara D. Martin and Dr. Antje Stoehr. Clara, I am thankful for all the support and the kindness that you provided me with throughout these years. I am honored that the big part of this thesis is a small part of your ERC project and I am looking forward to finally publishing something out of it… until now we have not been very lucky but hopefully, we will be in the future! Antje, I thank you for all the time, the meetings, the collaboration and the guidance you granted me during these years. Even though you were a novice in supervising PhD students, you showed straight away that you are very well suited for this role. You have been a referent during my PhD and I feel lucky for that. Vielen Dank auch, dass du mir geholfen hast, mein Deutsch zu bewahren!

Second, my gratitude goes to Dr. Anastastia Klimovich-Gray. Ana, you introduced the MEG world to me and I could not have been luckier. I started as a dumb toddler not even being able to run a MATLAB code and you made me a baby with few, yet solid, rudiments of programming. If this is not a miracle! I am also very grateful for your kindness, patience and support. I will preserve the memories of our moments together based on hard complaining and cursing (elegantly) things that were not working.

Third, I would like to thank Prof. Yury Shtyrov, Dr. Mads Jansen and Dr. Andreas Højlund at Aarhus University for our virtual collaboration. I am really looking forward to going deeper with our joint MEG project, part of which is presented in this thesis.

Fourth, I express my gratitude to the admin people, in particular Eider, Ana and Leire, the lab people, in particular Araitz, Manex, David, Mamen and Amets, and to the rest of the BCBL, especially the S&B Group, for giving me an enviable working context.

Fifth, I would like to thank La Caixa Foundation for supporting me, financially and morally, during my PhD. Without the INPhINIT program, I would have not been able to pursue the academic career and, especially, I would not have met the amazing people of my cohort. Lacaxitos, you rock!

Finally, I would like to thank my current colleagues at the University of Leipzig: Prof. Dr. Denisa Bordag and Dr. Andreas Opitz. Ich freue mich auf unsere Zusammenarbeit!

Acknowledgements (personal)

'It's very difficult todo esto' (M.R.). Now comes the part in which I express my affection and gratitude to the people that accompanied me throughout this bumpy journey. Normally one would say something like "I do not want to forget anyone, so I just say thank you to everybody" but I think that it simply means "I am too lazy to name all of you" and I do not want to do that. Yet, if I forget anyone, well sorry… You were important too! Warning: this is going to be a *chapa*, read at your own risk!

I start with the person with whom I started this crazy journey and with whom I shared amazing moments. Chiara, you are simply a wonderful human being: I do not know how I could describe you otherwise. Thank you for always being there, for cheering me up, for being a good friend. We are now far away from each other but I am sure that we will always find the way to meet and it will be as if we have never been apart.

Candice, this thesis would probably be inexistent without your friendship and support. Thank you for the sushi nights, the drinks and for listening my rants and whines. You are one of the most intelligent persons I have never met and I am grateful for having you as a friend.

Stefano, our years of 'flatmateship' were the best I have never had. I miss our Anoeta Palace so much! Thank you for being a generous and empathic friend. Thank you also for showing me that chaos is sometimes fine and that the world does not end if something is not perfectly clean. Finally, thanks for introducing me to your amazing mum, Paola, and your wonderful girlfriend, Merel.

Jose A. and Polina, thank you very much for being such lovely and supporting people. You folks are amazing and beautiful in all possible ways. I am looking forward to visiting you in Dublin!

Eneko, thanks for being a good friend who is always there no matter what. Asier, thank you for our conversations and for the, alas short, 'flatemateship'. Irene, you are one of the kindest persons I have never met. Thank you for having me at your place so many times and for the many moments we shared together. Dani, I do not think that I need to add anything to what I have already told you… just bear it in mind and thanks for everything. Vicente, Giorgio, Christoforos, Inés, Laura, Abraham, Jose P., Pierma, Jordi, Trisha, Jessie, Catherine, Romain, Florent, Mina, Carlos, Marta: having you in my everyday life was a blessing. Thank you for being part of the journey.

Special thanks to the amazing Italian community at the BCBL: Nicola, Nicoletta, Sendy, Simona and the already mentioned Chiara, Stefano and Pierma.

Bereziki eskerrik asko nire Euskara-irakasleari, Maialen. Zuri esker hizkuntz hau pixka bat hitz egin dezaket. Eredu bat zara! Quiero también dar un agradecimiento especial a Íñigo, gracias por todo, guapo!

I want to thank my friends who supported me regardless of the distance between us. Thank you Fede for being the best friend that one could ever wish. Ti voglio un bene dell'anima. Thank you and Fabio for giving me an 'acquired' niece, Celeste. My warmest thanks to Ale e Martina, Mikez, Rox and Arianna, Sara C. and Salvo, Sara Z., Ciccio, Ste, my buddy Luca, Giada e Teo, Alicc, Elena, Natalie, Fernando, Paola, Liza, Alessandro M., Alessandro R., Gabry, Martonzi, Friede and Olli (with Sonntag). You all make me feel so lucky, loved and supported.

Special thanks also to Dr. Marzia Minelli for her support during the darkest times and to Daria, whom I consider part of my family now.

Finally, I would like to thank my parents for their unconditional support for making me the person I am today.

Table of Contents

Appendices

Abstract

The present thesis presents a series of experiments that explore of the Orthographic Consistency Effect (OCE) in auditory language perception and production. The OCE is a psycholinguistic effect that shows a facilitation in processing words with sounds that can be spelled in only one way (e.g., /t/ in French or Spanish) in comparison to words with sounds that can be spelled in multiple ways (e.g., /k/ in French or Spanish). The OCE can be considered as a by-product of literacy since it is displayed by people who know how to read and write. In Experiment 1, the OCE in auditory language perception in L1-French and L1-Spanish adults was investigated by means of an auditory lexical decision task (LDT). Overall, the results showed that there was an OCE in only French pseudoword processing and Spanish word processing. The results were interpreted in light of different processing strategies related to the opacity of the writing systems (French is opaque, i.e., with many sound-to-spelling inconsistencies, while Spanish is transparent since it has only few sound-to-spelling inconsistencies). Experiment 2 investigated the OCE in Spanish seven-year-olds, who are at an early stage of reading acquisition, in auditory language perception using the same paradigm as in Experiment 1. The results showed that children show a strong OCE in both word and pseudoword auditory processing. The results were interpreted in light of developmental strategies in auditory language perception. Experiment 3 investigated the time course and the brain correlates of the OCE in auditory language perception. L1-Spanish adults were tested in an auditory LDT and a passive listening task. The results showed that the OCE occurs in an early time-window in auditory word processing in the LDT paradigm. As for the brain network involved, preliminary results showed that fronto-temporal regions are involved. The results were interpreted in light of brain models of auditory language processing. Experiment 4 explored the OCE in language production of French and Spanish adults by means of reading aloud and picture naming tasks. The results showed that orthography is not involved in language production. The findings were interpreted in the light of current models of reading and language production. Experiment 5 replicated Experiment 4 but involving Spanish seven-year-olds. The results showed an OCE only in reading aloud and they were interpreted in light of developmental accounts in reading

acquisition. Overall, this work showed that literacy can strategically influence language perception and production.

Resumen en castellano

Antecedentes teóricos

La presente tesis doctoral trata del impacto de la alfabetización, un fenómeno complejo. Para ello, este trabajo concibe la alfabetización como la capacidad de leer y escribir en idiomas con sistema de escritura fundamentado en el alfabeto latino, así como las capacidades de percepción auditiva y producción del lenguaje desde un punto de vista neurocognitivo. Por un lado, la *percepción auditiva del lenguaje* se refiere a los procesos cognitivos que transforman una señal acústica en abstracciones lingüísticas para concluir el acceso léxico (Poeppel, 2015). Por otro lado, la *producción del lenguaje* hace referencia a los procesos cognitivos que transforman una intención comunicativa en el habla (Dell & Jacobs, 2016).

Estos dos fenómenos psicolingüísticos se investigan a través de la observación de un efecto psicolingüístico llamado *efecto de consistencia ortográfica* (en adelante OCE, por sus siglas en inglés). Este efecto consiste en la mayor facilidad del procesamiento lingüístico como consecuencia de la relación unívoca entre unidades fonológicas y unidades ortográficas; en inglés, por ejemplo, la rima /əʊb/ se realiza ortográficamente solo con <-obe>, como en las palabras *globe* o *probe,* mientras la rima /eɪm/ se escribe <-aim> o <-ame>, como en las palabras *claim* or *flame*. De esta manera, se ha demostrado que el procesamiento de palabras como *globe* o *probe* resulta más rápido porque la rima activa solo una unidad ortográfica y, consecuentemente, el procesamiento lingüístico no está dificultado por la competición de otras formas ortográficas, como sería en el caso de *clame* o *flame* (Seidenberg & Tanenhaus, 1979).

Desde el punto de vista del desarrollo lingüístico, la alfabetización se produce a través de la *recodificación fonológica* (Ziegler & Goswami, 2005), que consiste en establecer unas conexiones entre las representaciones fonémicas existentes y sus correspondientes representaciones ortográficas todavía no adquiridas. Este proceso es diferente dependiendo de la opacidad del sistema de escritura de una lengua: lenguas “opacas” como el inglés o el francés presentan relaciones fonográficas muy inconsistentes (es decir, que múltiples grafemas corresponden a un mismo fonema) mientras lenguas “transparentes” como el español o el italiano presentan relaciones fonográficas bastante consistentes (esto es, que un único fonema corresponde a cada grafema). Según la Grain-Size Theory (Ziegler & Goswami), los alumnos que aprenden a leer una lengua opaca tienen que establecer relaciones fonográficas no solo a nivel fonológico, sino también a nivel silábico. Este fenómeno se debe a una estrategia de procesamiento del lenguaje que requiere unidades lingüísticas más complejas para realizar el acceso léxico de manera más eficiente, ya que el nivel fonémico es el más inconsistente de las relaciones fonográficas. Además, para los alumnos que aprenden una lengua transparente es suficientemente eficaz establecer relaciones fonográficas a nivel fonológico.

Experimento1

Teniendo en cuenta los enunciados de la Grain-Size Theory, la cuestión de si el OCE es un efecto que se presenta en los idiomas con sistemas alfabéticos con los mismos connotados durante la percepción auditiva del lenguaje se hace crucial. Más específicamente, queda todavía inexplorado si el OCE es un efecto que surge del nivel fonológico ya que estudios previos solo están basados en manipulaciones de los estímulos a nivel silábico (véanse los ejemplos en inglés mencionados anteriormente). Este aspecto ha sido considerado en Experimento 1 de la presente tesis.

Otro aspecto del que se ha tenido en cuenta en Experimento 1 es si un idioma come el castellano, notoriamente transparente y con muy pocas inconsistencias fonema-grafema, también presenta el OCE en la percepción auditiva del lenguaje. Ciertamente, los estudios hasta ahora publicados sobre este tema solo han investigado idiomas opacos como inglés (por ej. Dich, 2011) o francés (por ej. Ziegler & Ferrand, 1998) o un idioma más transparente pero con un rico inventario fonológico como el portugués (por ej. Ventura et al., 2004). El castellano representa un excelente candidato para la investigación del OCE en idiomas transparentes ya que, con la excepción del fonema /b/, todos los demás fonemas inconsistentes son sujetos a reglas ortográficas (por ej. /k/ se realiza con <c> antes de <o>, <u> o <a> mientras con <qu> con <e>, o <i>.

Finalmente, un último aspecto tratado en Experimento 1 es lo de las diferencias del OCE en idiomas en francés, un idioma opaco, y en castellano, un idioma transparente.

En breve, Experimento 1 consiste en una tarea de decisión léxica (LDT, desde el inglés *Lexical Decision Task*) en la que 30 estudiantes universitarios hablantes nativos de francés y 30 estudiantes universitarios hablantes nativos de castellano participaron. La tarea requiere que el participante escuche un estímulo, que puede ser una palabra o una pseudopalabra, y decidir lo más rápido y acertado posible si el estímulo es una palabra que existe en su idioma o no presionando una de las dos teclas a su disposición.

Las listas de los estímulos contienen 60 palabras y 60 pseudopalabras por cada idioma: la mitad de cada condición léxica es consistente y la otra inconsistente. Los estímulos consistentes contienen sólo fonemas consistentes (por ej. /t/ en francés y castellano) mientras los inconsistentes algunos fonemas inconsistentes (por ej. /k/ en francés y castellano). Es importante recalcar que el primer fonema de los estímulos inconsistentes es siempre una consonante inconsistente.

Las hipótesis son: a) el OCE es un efecto que se puede detectar también a nivel fonológico; b) un idioma transparente como el castellano también puede presentar un OCE; 3) el francés es el idioma que presenta un OCE más sólido ya que su sistema es más opaco que el castellano y la ortografía es un recurso más importante en la percepción auditiva del lenguaje para los francoparlantes. El OCE se traduce en tiempos de reacción (RT) más elevados para los estímulos inconsistentes respeto a los consistentes durante una LDT.

Los resultados revelan un OCE en la percepción auditiva de pseudopalabras en francés y en palabras en castellano: es decir, los francoparlantes son estadísticamente más rápidos en reconocer una pseudopalabra consistente que una pseudopalabra inconsistente mientras los castellanoparlantes son más rápidos en reconocer una palabra consistente que una palabra inconsistente. Respeto a las palabras franceses, se presenta solo una tendencia numérica que indica una facilitación durante el procesamiento de palabras consistentes mientras en las pseudopalabras castellanas el efecto es nulo. De un lado, los resultados del francés se interpretan como una evidencia de decodificación fonética durante el procesamiento de pseudopalabras que está afectado por las inconsistencias ortográficas de los fonemas. Este patrón no parece repetirse en el procesamiento de palabras en francés probablemente porque la unidad fonológica más determinante en la percepción auditiva de palabra en francés es la silaba y no el fonema y, por lo tanto, un sólido OCE se podría detectar solo si la manipulación de las palabras fuera a nivel silábico. Sin embargo, no es posible formular esta conclusión a la luz del Experimento 1 ya que este tema debe de ser afrontado en un estudio aparte. En relación al castellano, los resultados se interpretan basándose en la Grain-Size Theory ya que se presenta evidencia de decodificación fonológica durante la percepción auditiva de palabras. Sin embargo, los resultados de las pseudopalabras en castellano van en contra las expectativas ya que este tipo de estímulo debería ser procesado a nivel fonológico por definición puesto que no tienen representación léxica. No obstante, los resultados de las pseudopalabras están en consonancia con los resultados de estudios previos en portugués (Ventura et al., 2004), un idioma más cercano al castellano desde el punto de vista de la opacidad del sistema de escritura.

Experimento 2

Como mencionado el la introducción, aprender a leer implica la recodificación fonológica, es decir la creación de conexiones entre fonemas y grafemas. Esta premisa pone la cuestión de si niños en la fase de aprendizaje de la lectura pueden presentar un OCE en percepción auditiva del lenguaje puesto que sus relaciones fonográficas no están todavía consolidadas. El objetivo del Experimento 2 es investigar como los niños que aprenden francés y castellano están afectados por las inconsistencias de las relaciones fonema-grafema en sus respectivos idiomas nativos.

A causa de la pandemia de COVID-19, no ha sido posible recoger los datos de los niños franceses. Por ello, los participantes del Experimento 2 son 45 alumnos de siete años cuyo idioma nativo es el castellano. La tarea y los estímulos son exactamente los mismos del Experimento 1.

Los resultados revelan que los niños reconocen estímulos (es decir ambas palabras y pseudopalabras) consistentes más rápidamente y en manera más acertada que los inconsistentes. Respeto a los adultos castellanoparlantes, los niños que tienen un año y medio de instrucción a la lectura recurren masivamente a la decodificación fonológica durante la percepción auditiva del lenguaje. Este resultado se interpreta como una forma estratégica, dictada por la adquisición de la lectura, en la consolidación de las relaciones fonográficas y, probablemente, en la adquisición de vocabulario, aunque este último aspecto no puede ser afrontado con la LDT.

Experimento 3

Algunos investigadores afirman que el OCE es un efecto post-léxico en la percepción auditiva del lenguaje y estratégico en la realización de la tarea experimental demostrándolo con sus estudios empíricos (por ej. Cutler et al., 2010; Damian & Bowers, 2010). Sin embargo, una notable cantidad de estudios (por ej. Petrova et al., 2011) indican que el OCE no es un efecto estratégico sino generalizado de la percepción auditiva del lenguaje que requiere necesariamente la activación ortográfica, aunque el procesamiento sea en la modalidad auditiva, y que sobre todo el efecto ocurre a nivel pre-léxico, es decir antes que sobrevenga el acceso léxico. Sin embargo, las contribuciones científicas que más rigorosamente ponen a prueba estas dos hipótesis son los estudios de electroneurofisiología con electroencefalografía (EEG), ya que las medidas conductuales representan el "producto final" del proceso cognitivo que subyace la percepción auditiva del lenguaje (Perre & Ziegler, 2008). Desafortunadamente, los estudios que contribuyen a la discusión sobre el OCE pre- o post-léxico son pocos y principalmente con el francés como idioma de referencia.

Otro debate sobre el OCE abarca los circuitos cerebrales que subyacen la interrelación entre fonología y ortografía en la percepción auditiva del lenguaje. En breve, la literatura científica presenta dos hipótesis alternativas: a) la alfabetización causa una restructuración de las representaciones fonológicas y, por ello, esas contienen también información ortográfica. A nivel cerebral, el OCE se debería observar en zonas específicas del giro frontal inferior (GFI; por ej. la porción opercular) que, a pesar de estar relacionadas con el procesamiento fonológico, se han demostrado involucradas también en alguno procesos ortográficos en la percepción auditiva del lenguaje (Perre et al., 2009); b) la alfabetización construye un circuito cerebral que conecta las zonas típicamente relacionadas al procesamiento fonológico (GFI) y regiones cerebrales relacionadas al procesamiento ortográfico como el giro fusiforme. Eso implica que las representaciones fonológicas y las ortográficas son distintas y interactúan las unas con las otras (Bolger et al., 2008). Ahora bien, los estudios que presentan evidencia para una o la otra hipótesis se han llevado a cabo con técnicas de neuroimagen inadecuadas para le detección de un efecto que ocurre en los primeros 100ms desde la presentación del estímulo auditivo como la resonancia magnética funcional (Bolger et al., 2008) o con técnicas de localización de la señal electroneurofisiológica que presentan una resolución espacial muy limitada (Perre et al., 2009).

El Experimento 3 tiene entonces como objetivos lo de ampliar la base empírica sobre el debate del desarrollo temporal del OCE y, al mismo tiempo, de investigar los circuitos cerebrales que subyacen este efecto. Una técnica que presenta una buena resolución temporal y espacial es la magnetoencefalografía (MEG) y, por eso, ha sido adoptada en el experimento.

30 estudiantes castellanoparlantes nativos, los cuales no participaron a ningún otro experimento de esta tesis, han tomado parte en el Experimento 3 y han completado dos tareas mientras se registraba su actividad magnética cerebral a través de un escáner MEG con 306 sensores: una LDT y una tarea auditiva pasiva. La tarea auditiva pasiva se adopta para comprobar la omnipresencia del OCE en la percepción auditiva del lenguaje. Similarmente al Experimento 1, la lista de los estímulos contiene 50% palabras y pseudopalabras consistentes y 50% palabras y pseudopalabras inconsistentes, para un total de 320 estímulos, y son los mismos para las dos tareas.

Respeto a la LDT, los datos de 27 participantes (tres han sido excluidos por tener una señal muy ruidosa) se han analizado con un análisis espacio-temporal de los sensores MEG, en específico con un test de permutaciones para las palabras y las pseudopalabras separadamente. Los resultados de la LDT muestran una diferencia en la señal magnética de las palabras inconsistentes en la ventana temporal 56-150ms lo que indica que el OCE ocurre contemporáneamente al procesamiento fonológico y, entonces, antes del acceso léxico. Contrariamente a las expectativas, el patrón no se detecta en las pseudopalabras. Después de una observación detallada de los estímulos auditivos, se relata que la duración de la primera silaba de las pseudopalabras eras significativamente más larga que la de las palabras y, por lo tanto, no se puede excluir que esto prejuzga los resultados de las pseudopalabras. Los análisis de los sensores MEG muestran que una agrupación de sensores de la zona fronto-temporal izquierda ha detectado la diferencia de la señal magnética aportando una evidencia preliminar en favor de la hipótesis de restructuración fonológica.

Respeto a la tarea auditiva pasiva, los mismos análisis de la LDT no muestran ninguna diferencia en la señal entre estímulos consistentes o inconsistentes. Eso implica que probablemente haya una forma estratégica en el recurrir a la ortografía durante la percepción auditiva del lenguaje.

Experimentos 4 y 5

Los Experimentos 4 y 5 exploran si la alfabetización influye en la producción del lenguaje y en qué medida. El OCE en la producción del lenguaje no ha sido muy investigado, especialmente de manera interlingüística, y los resultados informados hasta ahora son contradictorios. El Experimento 4 consiste en una tarea de lectura en voz alta (RAT; desde el inglés Reading Aloud Task) y una tarea de denominación de imágenes (PNT; desde el inglés Picture Naming Task) en las que participaron los mismos participantes del Experimento 1. La manipulación de los estímulos es siempre a nivel fonético. En la RAT, tanto las palabras como las pseudopalabras pueden ser consistentes (es decir, con fonemas con un solo grafema correspondiente) o inconsistentes (es decir, con fonemas con múltiples grafemas correspondientes). En el PNT, los referentes de las imágenes eran consistentes o inconsistentes. La hipótesis es encontrar un OCE especialmente en la producción de pseudopalabras. Eso significa que se esperan tiempos de planificación y producción más breves para los estímulos consistentes que para los inconsistentes. Sin embargo, los resultados de RAT y PNT no informan ninguna diferencia significativa en la condición de consistencia.

El Experimento 5 consiste en las mismas tareas del Experimento 4 per en las que han participado los mismos participantes que el Experimento 2, es decir los alumnos de siete años castellanoparlantes. En breve, los resultados muestran que los niños españoles de siete años fueron más rápidos y precisos en la producción de estímulos (es decir palabras y pseudopalabras) consistentes en el RAT, pero no en el PNT.

Conclusiones

La presente tesis aporta unos resultados novedosos sobre la influencia de la alfabetización en la percepción auditiva del lenguaje. Por la primera vez, se demuestra que también idiomas transparentes como el castellano pueden presentar un efecto de consistencia ortográfica (OCE), ya que relaciones fonográficas consistentes causan una facilitación del procesamiento lingüístico en la modalidad auditiva. Desde el punto de vista del desarrollo lector, se expone que en la primera etapa de la adquisición de la lectura los niños recorren a las relaciones fonográficas tanto en la percepción auditiva del lenguaje como en la lectura en voz alta. Diversamente, los adultos no parecen recurrir a estas relaciones en la producción del lenguaje probablemente porque no las necesitan en este contexto.

Desde el punto de vista cerebral, la presente tesis relata sobre una investigación más sistemática y con una técnica de neuroimagen adecuada para el estudio del desarrollo temporal y de los circuitos cerebrales que subyacen el OCE.

Chapter 1 – Introduction

1.1 – Foreword

The present doctoral thesis addresses the topic of the influence of literacy in two psycholinguistic domains: auditory language perception and language production.

Throughout this work, the concept of *literacy* should be interpreted in its most narrow sense, namely the cognitive ability to read and write. This assumption is fundamental because the impact of literacy is investigated by looking at a specific psycholinguistic phenomenon called the *Orthographic Consistency Effect* which is abbreviated with OCE throughout this work. I describe the OCE in more detail later in this chapter. In brief, the OCE is the cognitive facilitation in both language perception and, to some extent, language production of words with consistent (i.e., with one-to-one) sound-to-spelling mappings. The cognitive facilitation is dictated by the fact that words with inconsistent (i.e., one-to-many) sound-to-spelling mappings are processed more slowly and less accurately due to the many orthographic representations of the correspondent phonological unit. One could claim that the OCE is an hindering effect because inconsistent sound-to-spelling mappings slow down auditory word perception. I argue that these two interpretations are two faces of the same coin and it is quite complicated to disentangle whether is one or the other: I used, thus, both accounts interchangeably throughout the thesis. In conclusion, the OCE is a clear by-product of literacy, because it originates from the relationship between sound and spelling. This relationship is indeed established during literacy acquisition (Malmstrom, 1975).

Auditory language perception is intended in this work as thus refers to the "set of operations that transform an auditory signal into representations of a form that makes contact with internally stored information—that is, the stored words in a listener's mental lexicon" (Poeppel, 2015, p. 429). Critically, some researchers—including Poeppel in the citation I have just reported—use the term *speech perception* to mean what I call *auditory language perception*. The reader should simply bear in mind that I adopted *auditory language perception* rather than *speech perception* because the latter is sometimes intended as the process of "mapping between properties of the acoustic signal and linguistic elements such as phonemes and distinctive features” (Diehl et al., 2004, p. 150). This definition of *speech perception* is limited to the phonetic and phonological level of auditory language perception which is not the only level of analysis that the present work considers.

Language production can be defined as a process of “determining the semantic content of one’s utterance (*conceptualization*), translating that content into linguistic form (*formulation*) and *articulation*” (Dell & Jacobs, 2016, p. 209). This means that I adopted this term to mean the cognitive processes underlying speaking. I reckon that *language production* is quite a broad term that could also imply other modalities (i.e., written or signed). Nevertheless, I will adhere to the most broadly used term in the field to describe speaking.

1.2 – Implications of being literate

Before presenting the relevant literature for this thesis, I would like to briefly present the different neurocognitive aspects that are related to literacy with the aim of contextualizing the work presented in this thesis.

1.2.1 – Literacy and the brain

There is a significant body of evidence that shows that literacy reshapes the brain by establishing and strengthening brain correlates involved in language processing and in cognition in general (see e.g., Dehaene et al., 2010). The striking impact of literacy on brain networks has been shown not only in children but also in adults (López-Barroso et al., 2020). This implies that a literate person conceives, processes, and utilizes language in a way that is strongly affected by the fact itself of being able to read and write. At the perceptual level, for example, the directionality of the writing system determines the dominance of the ear with which language is most perceived: this means that readers of scripts that go from left to right show a left-ear dominance in language perception while readers of scripts that go from right-to-left show the opposite pattern (Bertelson, 1972). Finally, concerning auditory language perception, literacy facilitates the recognition of spoken words and spoken language in general (for an overview see Morais & Kolinsky, 2019).

From this brief and surely inexhaustive report on the milestones in the literacy research, it becomes clear that the OCE is an effect that originates from literacy itself.

1.2.2 – Becoming literate in alphabetic languages

The present thesis focuses only on languages that use the Latin alphabet and any conclusion that I draw throughout my work refer to this type of languages. This does not imply, however, that there is no OCE in languages that have a writing system that is not alphabetic; there is, in fact, a growing body of evidence of the OCE in logographic languages (e.g., Chinese; Lee et al., 2015) but I will not cover this in the present work as the main focus is on French and Spanish—alphabetic Romance languages.

Phonological awareness, a metaphonological skill, is considered a predictor of reading acquisition in alphabetic languages (Tunmer & Rohl, 1991). Phonological awareness is the ability to access speech units and manipulate speech sounds. For example, preschoolers that are taught to read through rime analogy (i.e., reading words that rhyme with each other) and other prereading skills (e.g., initial phoneme identity) acquire reading more easily (Walton & Walton, 2002). The first step in reading acquisition is *phonological recoding* which means that learners map phonological representations onto orthographic representations (Ziegler & Goswami, 2005). In other words, children (or adults) match already existing phonemic representations with new orthographic representations (cf. Morais, 2021).

The importance of the Grain Size Theory by Ziegler and Goswami (2005) lies in recognizing developmental differences in reading acquisition given that alphabetic languages vary in the consistency of their sound-to-spelling mappings. Grain Size Theory posits three aspects that are crucial in reading acquisition: *availability*, *consistency,* and *granularity*. The first refers to accessibility of the different phonological units (e.g., phoneme, syllables, body, or rime) before reading acquisition. The second comprises both the opacity of the writing system but also the degree of inconsistency of the different phonological units with the respective orthographic representations (i.e., phoneme-to-grapheme mappings or bigger units to the respective spellings). The latter refers to the fact that there are larger linguistic units (e.g., words) than basic linguistic units (e.g., graphemes) and they need to be learnt based on the grain size of the critical phonological units in a given language. For example, learners of French will need to integrate the orthographic representations of both syllables and phonemes because knowing the phoneme-to-grapheme correspondences is not enough to establish spelling. There are syllables-to-orthography correspondences that are both more informative and more consistent and therefore need to be learned directly in parallel. Conversely, learners of Spanish usually only acquire phoneme-to-grapheme mappings because their language is quite transparent, and most phonemes have just one corresponding grapheme, regardless of the syllable it is contained in. The Grain Size Theory is particularly relevant to this thesis because it is the cornerstone on which the stimulus manipulation is based in all experiments. More details follow in Chapter 2.

1.3 – The OCE in language perception

As I briefly reported in the previous paragraph, literacy influences auditory language perception and one of the most relevant pieces of evidence is the OCE. From my point of view, the OCE is particularly interesting because it shows that being literate could be "disadvantageous" since an auditory linguistic stimulus with inconsistent sound-to-spelling mappings hinders auditory language perception. Since two thirds of this thesis cover the OCE in auditory language perception, this section presents a detailed literature review of the OCE in this domain.

1.3.1 – Consistency effects in visual and auditory modality

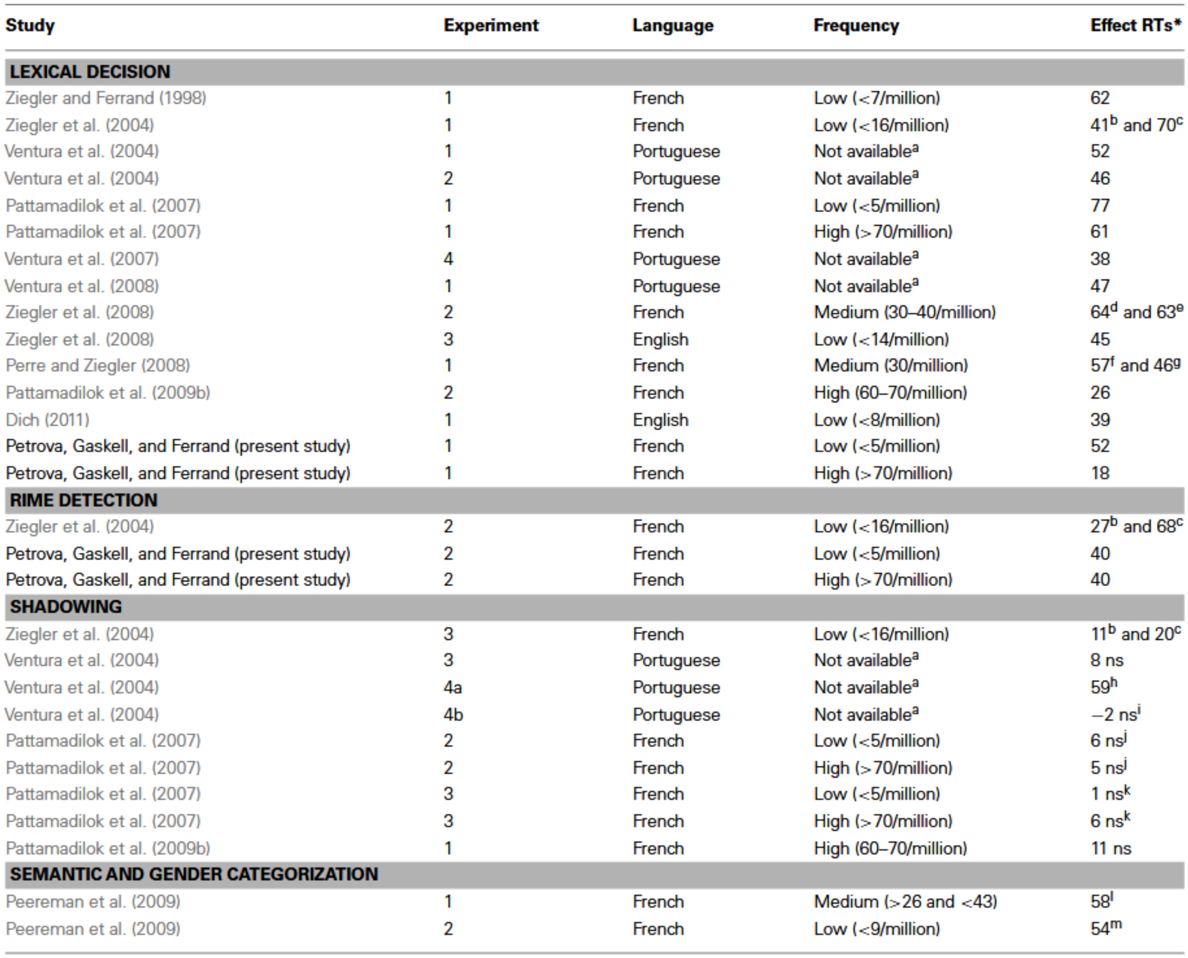
First, it should be noted that the research on the OCE in the auditory modality was inspired by investigations in the visual modality (e.g., D. E. Meyer et al., 1974). In the framework of the visual modality, the effect of (in)consistent sound-to-spelling mappings on visual word processing is often referred as *feedback consistency effect* while spelling-to-sound (in)consistencies cause the *feedforward consistency effect* (e.g., Stone et al., 1997). Similarly, studies in the field of auditory language perception (e.g., Frost & Katz, 1989; Ziegler & Ferrand, 1998) show that there are *feedforward* and *feedback* consistency effects in the auditory domain. The feedback consistency effect in the auditory modality has also been labeled as Orthographic Consistency Effect (Seidenberg & Tanenhaus, 1979; Ventura et al., 2004). I chose to use this term throughout the thesis for the sake of clarity.

1.3.2 – The OCE in the literature

Seidenberg and Tanenhaus (1979) is often cited as the first contribution to the literature about the OCE, where this effect was found in a rhyme detection task. In brief, the results showed that participants were faster at determining that pairs of words with consistently-spelled rimes (e.g., *pie* and *tie*) rhymed compared to pairs of words with inconsistently-spelled rimes (e.g., *rye* and *tie*). According to the authors, an important implication of these results is that orthography is automatically accessed in auditory word processing. Yet, up until Ziegler and Ferrand (1998), all studies on the OCE were based on tasks where the effect was prompted by primes as in rhyme detection tasks (e.g., Seidenberg & Tanenhaus, 1979), primed auditory lexical decision tasks (e.g., Jakimik et al., 1985), and phoneme monitoring tasks (e.g., Frauenfelder et al., 1990). Therefore, Ziegler and Ferrand (1998) were the first to test whether the OCE emerges in a more ecological paradigms such as yes/no auditory lexical decision tasks (LDT). Overall, the results showed an OCE for word recognition in which participants were quicker and more accurate in recognizing consistent than inconsistent French words. No OCE was reported, however, for pseudowords. The authors motivated the absence of the OCE in auditory pseudoword perception with the argument that pseudowords might undergo different linguistic processes. A possible explanation that Ziegler and Ferrand gave for the absence of the OCE in auditory pseudoword perception is based on the "timing-out" mechanism. This mechanism posits that participant of an LDT respond "no" to pseudowords once the threshold for lexical activation is not reached (i.e., no correspondent lexical entry has been found) and this cancels out any possible OCE. Yet, an OCE in auditory pseudoword perception was found in later studies with French as a target language and with an LDT as experimental paradigm (Pattamadilok et al., 2007a; Pattamadilok, Perre, et al., 2009). The OCE in auditory pseudoword perception is still an open debate. As Taft (2011) argued, some studies did not have very well matched stimuli and some others report larger effect sizes than other with very similar characteristics. I argue that the investigation of the OCE at the phonemic level better informs this debate because pseudowords are usually decoded in auditory language perception and phonological decoding occurs at the phonemic level.

Despite the contradicting findings on the OCE in auditory pseudoword perception, the OCE in auditory word perception has been reliably found in many studies and through different experimental designs, as Table 1 summarizes.

Table 1. Summary of the behavioral evidence of the OCE divided by task type.



Study	Experiment	Language	Frequency	Effect RTs*
LEXICAL DECISION				
Ziegler and Ferrand (1998)	1	French	Low (<7/million)	62
Ziegler et al. (2004)	1	French	Low (<16/million)	41[b] and 70[c]
Ventura et al. (2004)	1	Portuguese	Not available[a]	52
Ventura et al. (2004)	2	Portuguese	Not available[a]	46
Pattamadilok et al. (2007)	1	French	Low (<5/million)	77
Pattamadilok et al. (2007)	1	French	High (>70/million)	61
Ventura et al. (2007)	4	Portuguese	Not available[a]	38
Ventura et al. (2008)	1	Portuguese	Not available[a]	47
Ziegler et al. (2008)	2	French	Medium (30–40/million)	64[d] and 63[e]
Ziegler et al. (2008)	3	English	Low (<14/million)	45
Perre and Ziegler (2008)	1	French	Medium (30/million)	57[f] and 46[g]
Pattamadilok et al. (2009b)	2	French	High (60–70/million)	26
Dich (2011)	1	English	Low (<8/million)	39
Petrova, Gaskell, and Ferrand (present study)	1	French	Low (<5/million)	52
Petrova, Gaskell, and Ferrand (present study)	1	French	High (>70/million)	18
RIME DETECTION				
Ziegler et al. (2004)	2	French	Low (<16/million)	27[b] and 68[c]
Petrova, Gaskell, and Ferrand (present study)	2	French	Low (<5/million)	40
Petrova, Gaskell, and Ferrand (present study)	2	French	High (>70/million)	40
SHADOWING				
Ziegler et al. (2004)	3	French	Low (<16/million)	11[b] and 20[c]
Ventura et al. (2004)	3	Portuguese	Not available[a]	8 ns
Ventura et al. (2004)	4a	Portuguese	Not available[a]	59[h]
Ventura et al. (2004)	4b	Portuguese	Not available[a]	−2 ns[i]
Pattamadilok et al. (2007)	2	French	Low (<5/million)	6 ns[j]
Pattamadilok et al. (2007)	2	French	High (>70/million)	5 ns[j]
Pattamadilok et al. (2007)	3	French	Low (<5/million)	1 ns[k]
Pattamadilok et al. (2007)	3	French	High (>70/million)	6 ns[k]
Pattamadilok et al. (2009b)	1	French	High (60–70/million)	11 ns
SEMANTIC AND GENDER CATEGORIZATION				
Peereman et al. (2009)	1	French	Medium (>26 and <43)	58[l]
Peereman et al. (2009)	2	French	Low (<9/million)	54[m]

Note: ‘Frequency’ refers to the word frequency of the stimuli. ‘Effect RT’ refers to the magnitude of the OCE in milliseconds (retrieved from Petrova et al., 2011, p. 3).

This overview shows that the OCE is a robust effect in the auditory modality, and it also appears not to depend on word frequency. Petrova, Gaskell, and Ferrand (2011) orthogonally manipulated word frequency and found the OCE in all conditions and in both lexical decision task and rime detection.

Another issue related to the OCE is the highly debated question of whether it is an online, pre/sub-lexical or an offline, (post-)lexical effect. Ventura and colleagues (2004) made a strong argument in favor of the post-lexicality of the OCE. They claimed that the OCE requires lexical access in order to occur because they did not find any OCE in shadowing task (i.e., a task that elicits online processes) but only in lexical decision task with Portuguese, a fairly transparent language. However, it should be noted that it is difficult to disentangle the perception and production processes in the shadowing (Cutler, 1995). In some studies employing a rime detection paradigm, the authors argued that the OCE is a strategic effect in which listeners might resort to orthography when completing an experimental task (Damian & Bowers, 2010; Ziegler et al., 2004). Conversely, Petrova and colleagues (2011) believe that the OCE is an online process because rime detection tasks require necessarily word segmentation. Moreover, they also found no significant difference in the magnitude of the OCE between the rime detection and the lexical decision task, in contrast to Ziegler and colleagues (2004).

1.3.3 – Modeling approaches to the OCE

From a modeling perspective, there has been no dedicated investigation on the OCE in the auditory modality. On the one hand, models for auditory language perception do not include any interaction with orthography (e.g., the TRACE model—McClelland & Elman, 1986—or the Neighbor Activation Model—NAM; Luce & Pisoni, 1998). All the models that take into account the phonology-orthography interface were proposed for reading or reading aloud (e.g., the dual-route cascade model—DRC; Coltheart & Rastle, 1994—, the Connectionist Dual-Process model—CDP+; Perry et al., 2007, 2010, 2013—, or the bimodal interactive activation model BIA; Grainger & Ferrand, 1996). According to Frost and Ziegler (2012), a model that is potentially compatible with the feedback consistency effects and the OCE in both visual and auditory modality is the BIA (see Figure 1). The most important feature of the BIA model is the symmetric structure across the visual and auditory modality. This symmetry makes the OCE in the auditory modality a mirror of the feedback consistency effect in the visual modality and vice versa. The BIA model posits that an auditory input activates the sub-lexical phonological representations which, in turn, activate the correspondent orthographic representations at the same level. This interaction, together with the interaction between the sub-lexical and the lexical representations, eventually leads to lexical access (i.e., word recognition). The BIA model does not explicitly posit the different sub-lexical units (i.e., phonological, syllabic, body, or coda).

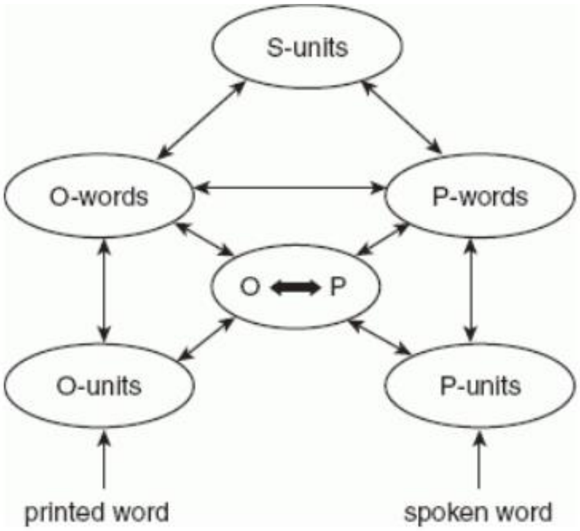


Figure 1. Schematic representation of the BIA model (retrived from Frost & Ziegler, 2012)

Nevertheless, the BIA model could easily take into account these different units depending on the opacity of the language as Jacobs and colleagues simulated in an extension of the BIA model (Jacobs et al., 1998). In Experiment 1, the issue of the phonological sub-lexical units is addressed in more detail. For now, I would like to point out that, despite the large body of research on the OCE in auditory language perception (see Table 1), the alphabetic languages investigated in this framework are few and mostly opaque. This represents an important gap in the literature also considering that none of the abovementioned models has simulated the OCE.

1.4 – The OCE in the neurolinguistic literature

There is a body of research that investigates the OCE in auditory language perception using neuroimaging techniques. In the present section, I illustrate the debate on two important issues in the neurolinguistic literature related to the OCE: the time frame and the brain correlates of the OCE in this modality.

1.4.1 – The time frame of the OCE

As I will argue in more detail in Chapter 3, the pre-lexical vs. post-lexical debate of the OCE cannot be solved with behavioral studies. Behavioral measures like response times (RT) or accuracy do not inform us on the temporal development of the effect. This aspect can only be measured by electroneurophysiological techniques like M/EEG. Perre and Ziegler (2008) were the first to address the issue of whether the OCE is an online⁄pre-lexical effect or post-lexical⁄decisional artifact by using EEG in auditory LDT. The authors expected that a hypothetically post-lexical⁄decisional OCE should affect late ERP components such as the late positive component (LPC). Conversely, a hypothetically online⁄pre-lexical OCE should modulate ERP components related to the N400, known for being implicated in lexical access. Moreover, the authors manipulated their inconsistent stimuli to get an early inconsistency (i.e., initial vowel cluster with multiple possible spellings like ⁄ry⁄ in French) or a late inconsistency (i.e., final vowel cluster with multiple possible spellings in ⁄os⁄). Behaviorally, inconsistent words were recognized more slowly and less accurately than consistent words. Yet, no difference between early and late inconsistency was found. As for the EEG results, the OCE was localized over centro-posterior electrodes due to an interaction between brain region and consistency. From this electrode site, the earliest amplitude difference between early inconsistent words and consistent words emerged in the 300−350ms after word onset with an enhanced negativity in the early inconsistent condition. Similarly, late inconsistent words generated a larger negativity in the 350−700ms time window. The authors argued that the most striking result is that the negativity peak emerged around 160−190ms after the onset of the inconsistent chunk of the stimulus. Overall, Perre and Ziegler claimed that their results show that the OCE is an online and pre-lexical effect. Moreover, the N320 component found in this study, which refers to the early OCE, is related to the sub-lexical activation of phonology from printed words (Bentin et al., 1999).

A similar pattern was found by Pattamadilok and colleagues (2009) in a semantic go⁄no-go task that provided stronger evidence of the pre-lexicality of the OCE. In another study, Pattamadilok, Perre, and Ziegler (2011) investigated whether the neural correlates of the OCE emerging from a metaphonological task (i.e., rime detection) can be related to those from a lexical decision. In contrast to non-metaphonological tasks, the results of this study showed an OCE in the P200 time frame, where inconsistent words elicited an enhanced signal. Moreover, a late consistency effect (yet with an inversed pattern, i.e. higher amplitude for the consistent stimuli) emerged in the 350−700ms time window, coinciding with the decisional process of the task (i.e., saying “yes” if the two words rhymed or “no” if they did not). In sum, the authors claim that the OCE can be task-specific. The early P200 effect, for example, can be related to a task-specific phonological process linked to segmentation (a psycholinguistic process needed in this kind of task) which is not needed in a lexical decision task. The late negativity effect shows that the OCE modulates the decisional process required in the task and not vice versa.

1.4.2 – The neural correlates of the OCE

Another important issue that the neurolinguistic literature addresses is related to the localization of the neural network underlying the OCE in auditory language perception. Parallel to the models presented in the previous section, there are two viewpoints on the matter (see Figure 2). The first posits that the there is a co-activation of both orthographic and phonemic representations (e.g., BIA model; Grainger & Ferrand, 1996). The second point of view proposes that phonemic representations are restructured by literacy (e.g., Goswami et al., 2005). In other words, orthographic information is embedded in phonemic representations: if a phonological unit has multiple spellings, then the phonemic representation of that sound can be seen as more "dispersed" and/or with a higher threshold for its activation. Conversely, if a phoneme has just one way to be spelled, then its representation is more "condensed" or has a lower activation threshold (Pulvermüller, 1999).

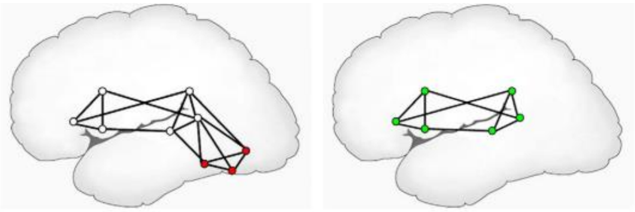


Figure 2. Adaptation of the Pulvermüller's language brain network accounts representing the two hypotheses: the left panel illustrates the orthographic co-activation hypothesis and the right panel the phonemic restructuring (retrieved from Perre et al., 2009).

Montant and colleagues (2011) tested the recalibration and the co-activations accounts with an fMRI study involving an auditory lexical decision task and a related auditory control task in which participants had to determine whether the first of a series of pronounced vowels was produced by a male or a female speaker. If the co-activation theory is true, an activation of both the VWFA and the spoken language neural system (i.e., inferior temporal gyrus—IFG; superior temporal gyrus—STG; supramarginal gyrus—SMG) should occur. On the contrary, if there is a phonological restructuring due to literacy, there would not be any activation in those brain areas devoted to visual word processing. The results showed an ample activation of the IFG, the anterior insula, the STG, and the middle temporal gyrus in the control task. As for the consistent-inconsistent contrast, the blood-oxygen-level-dependent (BOLD) response displayed greater activation of the IFG for inconsistent words. Therefore, the authors claimed that their findings support the phonemic restructuring argument given that there was no involvement of brain areas related to orthographic processing.

These results relate to two previous studies also using auditory lexical decision tasks: one employing transcranial magnetic stimulation (TMS; Pattamadilok et al., 2010) and the other conducted with EEG and sLORETA source reconstruction (Perre et al., 2009).

In short, the TMS study showed that the OCE vanished only when the stimulation was directed to the left SMG but not when it was operated on the left ventral occipito-temporal cortex (vOTC). That is, the reaction times (RT) of consistent words increased when the TMS was operated on the SMG and there was no difference with the RTs of inconsistent words. Yet, when the vOTC was stimulated, consistent words were still recognized faster than inconsistent words. These findings suggest that there is no co-activation of orthographic representations related but rather a restructuring of phonemic representations. However, the disappearance of the OCE due to the inhibition of the phonological areas might be the result of other disruptive factors deriving from the TMS itself (e.g., blocking the interaction with the VWFA). As for the non-disruption of the OCE with the TMS on the VWFA, it could be interpreted that that brain area is not strictly necessary for lexical access but yet it does not imply that it is not involved at all in that process.

As for the EEG study with sLORETA source reconstruction, the researchers measured the EEG signal during an auditory lexical decision task but, contrary to Perre and Ziegler (2008), the stimuli were manipulated only in the first syllable (as in the early inconsistency condition in Perre and Ziegler, 2008). As in previous EEG studies (Pattamadilok et al., 2009; Perre & Ziegler, 2008), a larger negativity was detected in the 300−380ms and 410−550ms time windows for inconsistent words. Concerning the source reconstruction with sLORETA, a significant difference in the activation of the left temporo-parietal area was found in the first time-window related to the consistency condition. No difference was found in the VWFA and in the second time window. This findings support, therefore, the restructuring account as well.

It should be noted that the studies mentioned up to now were all conducted in French and, most interestingly, the theoretical implications clash with those of other studies in the literature. An fMRI study (Booth et al., 2007) involving children aged 7 to 15 years performing an auditory lexical decision task in English revealed a co-activation of the IFG and surrounding areas related to phonological processing but also the VWFA which supports the co-activation account. More specifically, an age-related increase of activation of the left inferior parietal region for the words with inconsistent rimes suggests that more skilled readers can rely on more advanced sound-to-spelling mappings. In fact, these results are in line with another fMRI study with adults performing a rime detection task in English (Booth et al., 2003).

To conclude, an EEG study with source reconstruction sLORETA (Chen et al., 2016) challenged Perre and colleagues' (2009) findings with a go/no-go semantic task in Chinese. Chinese characters are usually composed of two parts, called radicals. One radical is defined as phonetic since it provides information on the pronunciation of the word. The other radical is usually defined as semantic since it provides information on the meaning of the word. Therefore, the stimulus manipulation that basically mirrors what has been done with alphabetic languages with the (in)consistent sound-to-spelling mappings. Even though no significant results related to the OCE were found behaviorally, the OCE was detected in the N400 time window. It is beyond the scope of this review to discuss in detail the results of this study. Yet, it is worth noting that the sLORETA source reconstruction revealed that the OCE resulted in a different activation not only of frontal and temporo-parietal areas but also of VWFA which, in contrast to Perre and colleagues (2009), supports the co-activation account.

In sum, the literature presents empirical evidence for both accounts. I argue in Chapter 3, however, that none of the previous studies adopted a neuroimaging technique that could reliably capture both the time course and the brain networks involved in auditory language perception.

1.5 – The influence of literacy in language production

A classic and straightforward example of how literacy can influence language production is the study by Stroop (1935). The Stroop test, which originated from the 1935 study, is one of the most famous tasks in cognitive psychology in which a participant has to name the ink color of a printed color name (e.g., saying red to the following stimulus: yellow). As simple as it seems, the task is particularly challenging for literate people because reading is a cognitive ability that cannot easily be inhibited and, therefore, it is difficult not to read the color name aloud and name the ink color instead.

In this section, I briefly address three issues related to the influence of literacy on language production: a) the role of literacy in reading aloud and picture naming paradigms, b) models on language production that consider phonology-orthography relations, and c) the perception-production link.

Excluding the difference between picture naming and reading aloud in the initial cognitive steps involved, these two paradigms share word-planning processes like, for example, phonological encoding (Araújo et al., 2019; Roelofs, 2004). This raises the question of whether phonology-orthographic relations are similarly, if at all, involved in phonological encoding during language production in the two paradigms. Roelofs (2006) has been the only one who systematically compared orthographic effects in the two paradigms. He found an OCE (i.e., a feedback consistency effect) only in reading aloud and concluded that the effect occurred because of a strategic use of sound-to-grapheme mappings in reading aloud related to the task itself.

I have already mentioned in section 1.3.3 a couple of models of language production that regard phonology-orthography relations which relate to reading aloud. It should be noted that models of language production can be classified into two families: the modular and the interactive. Modular models (e.g., Levelt et al., 1999) consist of well-defined cognitive processing stages which are organized hierarchically and occur in a sequential fashion (see Figure 3 – left). This implies that when a speaker utters a word, their non-linguistic, conceptual system first activates the related lexical-semantic unit (i.e., lemma) which in turn retrieves the corresponding lexical-phonological codes. The most important feature is that the flow of information is unidirectional and that the next step does not start if the previous one is not completed. This typology of models is particularly popular in the picture naming literature (see e.g., Glaser, 1992). However, the nature of these models does not allow a bidirectional flow of information between the phonological and orthographic levels (Damian & Bowers, 2010). For this reason, the few studies on the orthographic effects in language production always relate their findings with interactive models which, contrarily to modular models, posit bidirectional flows of information between modalities (i.e., auditory and visual) and between linguistic levels.

Since no computational model simulations are presented in this thesis, I will limit myself to the presentation of the dual-route cascade model (DRC), an interactive model, by Coltheart and Rastle (1994; further developed in Coltheart et al., 2001a). The choice of this model is based only on its theoretical assumptions and also because of the fact that it was specifically created for reading aloud (i.e., language production). A basic theoretical assumption that makes this model relevant to the study of the OCE in language production is the bidirectionality of the flow of information between the different levels (see Figure 3 – right). The DRC model is composed by three routes (the 1994 version had only two and from this came the label 'dual-route'): the lexical-semantic, the non-lexical-semantic and the grapheme-phoneme conversion (GPC) routes. The different layers of each unit interact with each other through excitation or inhibition. The first refers to the activation of one unit which facilitates the activation of the other units (e.g., a graphemic unit facilitates the activation of the phonemic unit). The second refers, conversely, to the activation of one unit which hinders the activation of the other unit. From the DRC model perspective, the OCE occurs in the non-lexical-semantic route in which orthographic and phonological level are involved and interact with each other. I refer the reader to the cited articles for more technical details about the DRC model.

Finally, I would like to point out that the study of the OCE in language production fits in the controversial scientific debate about the perception-production association. In a nutshell, the debate is based on the question of whether language perception and language production are two sides of the same coin because they are cooperative processes. The debate is particularly popular in the phonology literature which presents empirical evidence for both the association of the two domains (e.g., Nielsen, 2011) or the dissociation of the two (Baese-Berk & Samuel, 2016). No studies directly address this debate from the OCE perspective. Consequently, the present work represents a first attempt.

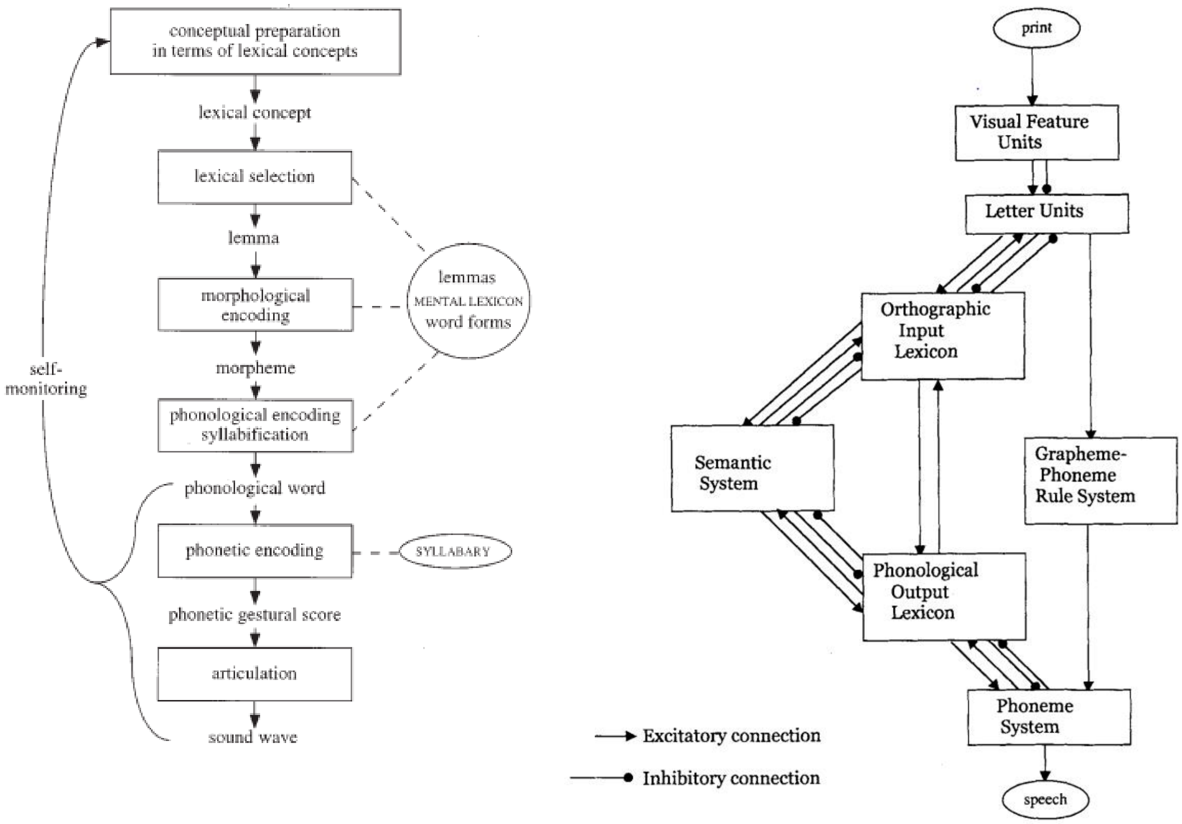


Figure 3. Right: an example of modular model of speech production (retrieved from Levelt et al., 1999). Left: an example of an interactive model (DRC model; retrieved from Coltheart et al., 2001a).

1.6 – Current Work

To investigate the OCE in auditory language perception and language production, I adopted two experimental approaches: cross-linguistic and developmental. The first was motivated by the fact that languages vary in the consistency of their phoneme-to-grapheme mappings and, therefore, the OCE could emerge to different extents in an opaque language compared to a transparent language. The second approach was motivated by the fact that reading acquisition shapes the way literates process language. Children at an early stage of reading acquisition are developing important reading skills such as decoding that will help them become expert readers. Investigating the OCE in this specific developmental stage is pivotal to understanding the effect itself. Little is known about when listeners start showing an OCE in auditory language perception.

In the following three chapters, I present experimental data that cover issues related to the OCE in auditory language perception and production, more specifically:

- In Chapter 2, Experiment 1 addresses the issue of whether there is an OCE in a transparent language like Spanish and whether it stems from the manipulation of phoneme-to-grapheme mappings in both French and Spanish. The choice of these two languages was aimed at comparing an opaque and a transparent language to explore possible cross-linguistic differences. In Experiment 2, I tested Spanish-speaking seven-year-olds to investigate whether they show an OCE even though they are still consolidating their phoneme-to-grapheme mappings through reading instruction. In both experiments, participants carried out auditory lexical decision tasks (LDT).
- In Chapter 3, Experiment 3 investigates the time course and brain networks underlying the OCE in auditory language perception by means of MEG. As I already mentioned, I specifically addressed the pre-lexical vs. post-lexical debate on the OCE and the co-activation versus restructuring accounts. The experimental tasks were an auditory LDT and a passive listening task. The participants were Spanish-speaking adults.
- In Chapter 4, I investigated the OCE in language production of French-speaking and Spanish-speaking adults (Experiment 4) and of Spanish-speaking children (Experiment 5) by employing both reading aloud and picture naming tasks.

Finally, Chapter 5 presents a general discussion in which I bring together the theoretical implications deriving from the findings of the experimental chapters.

References

Araújo, S., Fernandes, T., & Huettig, F. (2019). Learning to read facilitates the retrieval of phonological representations in rapid automatized naming: Evidence from unschooled illiterate, ex-illiterate, and schooled literate adults. *Developmental Science*, *22*(4), 1–9. https://doi.org/10.1111/desc.12783

Baese-Berk, M. M., & Samuel, A. G. (2016). Listeners beware: Speech production may be bad for learning speech sounds. *Journal of Memory and Language*, *89*, 23–36. https://doi.org/10.1016/j.jml.2015.10.008

Bentin, S., Mouchetant-Rostaing, Y., Giard, M. H., Echallier, J. F., & Pernier, J. (1999). ERP manifestations of processing printed words at different psycholinguistic levels: Time course and scalp distribution. *Journal of Cognitive Neuroscience*, *11*(3), 235–260. https://doi.org/10.1162/089892999563373

Bertelson, P. (1972). Listening from Left to Right versus Right to Left. *Perception*, *1*(2), 161–165. https://doi.org/10.1068/p010161

Booth, J. R., Burman, D. D., Meyer, J. R., Gitelman, D. R., Parrish, T. B., & Mesulam, M. M. (2003). *Relation Between Brain Activation and Lexical Performance*. *169*, 155–169. https://doi.org/10.1002/hbm.10111

Booth, J. R., Cho, S., Burman, D. D., & Bitan, T. (2007). Neural correlates of mapping from phonology to orthography in children performing an auditory spelling task. *Developmental Science*, *10*(4), 441–451. https://doi.org/10.1111/j.1467-7687.2007.00598.x

Chen, W. F., Chao, P. C., Chang, Y. N., Hsu, C. H., & Lee, C. Y. (2016). Effects of orthographic consistency and homophone density on Chinese spoken word recognition. *Brain and Language*, *157–158*, 51–62. https://doi.org/10.1016/j.bandl.2016.04.005

Coltheart, M., & Rastle, K. (1994). Serial processing in reading aloud: Evidence for dual-route models of reading. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(6), 1197–1211. https://doi.org/10.1037/0096-1523.20.6.1197

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204–256. https://doi.org/10.1037/0033-295X.108.1.204

Cutler, A. (1995). Chapter 4—Spoken Word Recognition and Production. In J. L. Miller & P. D. Eimas (Eds.), *Speech, Language, and Communication (Second Edition)* (pp. 97–136). Academic Press. https://doi.org/10.1016/B978-012497770-9.50006-6

Damian, M. F., & Bowers, J. S. (2010). Orthographic effects in rhyme monitoring tasks: Are they automatic? *European Journal of Cognitive Psychology*, *22*(1), 106–116. https://doi.org/10.1080/09541440902734263

Dehaene, S., Pegado, F., Braga, L. W., Ventura, P., Filho, G. N., Jobert, A., Dehaene-Lambertz, G., Kolinsky, R., Morais, J., & Cohen, L. (2010). How Learning to Read Changes the Cortical Networks for Vision and Language. *Science*, *330*(6009), 1359–1364. https://doi.org/10.1126/science.1194140

Dell, G. S., & Jacobs, C. L. (2016). Successful Speaking. In *Neurobiology of Language* (pp. 209–219). Elsevier. https://doi.org/10.1016/B978-0-12-407794-2.00018-3

Diehl, R. L., Lotto, A. J., & Holt, L. L. (2004). Speech Perception. *Annual Review of Psychology*, *55*(1), 149–179. https://doi.org/10.1146/annurev.psych.55.090902.142028

Frauenfelder, U. H., Segui, J., & Dijkstra, T. (1990). Lexical Effects in Phonemic Processing: Facilitatory or Inhibitory? *Journal of Experimental Psychology: Human Perception and Performance*, *16*(1), 77–91. https://doi.org/10.1037/0096-1523.16.1.77

Frost, R., & Katz, L. (1989). Orthographic depth and the interaction of visual and auditory processing in word recognition. *Memory & Cognition*, *17*(3), 302–310. https://doi.org/10.3758/BF03198468

Frost, R., & Ziegler, J. C. (2012). Speech and spelling interaction: The interdependence of visual and auditory word recognition. In *The Oxford Handbook of Psycholinguistics* (pp. 1–21). https://doi.org/10.1093/oxfordhb/9780198568971.013.0007

Glaser, W. R. (1992). Picture naming. *Cognition*, *42*(1–3), 61–105. https://doi.org/10.1016/0010-0277(92)90040-O

Goswami, U., Ziegler, J. C., & Richardson, U. (2005). The effects of spelling consistency on phonological awareness: A comparison of English and German. *Journal of Experimental Child Psychology*, *92*(4), 345–365. https://doi.org/10.1016/j.jecp.2005.06.002

Grainger, J., & Ferrand, L. (1996). Masked orthographic and phonological priming in visual word recognition and naming: Cross-task comparisons. *Journal of Memory and Language*, *35*(5), 623–647. https://doi.org/10.1006/jmla.1996.0033

Jacobs, A., Rey, A., Ziegler, J., & Grainger, J. (1998). MROM-P: An interactive activation, multiple read-out model of orthographic and phonological processes in visual word recognition. In J. Grainger & A. Jacobs (Eds.), *Localist Connectionist Approaches to Human Cognition* (pp. 147–188). Lawrence Erlbaum Associates Publishers.

Jakimik, J., Cole, R. A., & Rudnicky, A. I. (1985). Sound and spelling in spoken word recognition. *Journal of Memory and Language*, *24*(2), 165–178. https://doi.org/10.1016/0749-596X(85)90022-1

Lee, C.-Y., Hsu, C.-H., Chang, Y.-N., Chen, W.-F., & Chao, P.-C. (2015). The Feedback Consistency Effect in Chinese Character Recognition: Evidence from a Psycholinguistic Norm. *Language and Linguistics*, *16*(4), 535–554. https://doi.org/10.1177/1606822X15583238

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 38–75. https://doi.org/10.1017/s0140525x99001776

López-Barroso, D., Thiebaut de Schotten, M., Morais, J., Kolinsky, R., Braga, L. W., Guerreiro-Tauil, A., Dehaene, S., & Cohen, L. (2020). Impact of literacy on the functional connectivity of vision and language related networks. *NeuroImage*, *213*, 1–12. https://doi.org/10.1016/j.neuroimage.2020.116722

Luce, P. A., & Pisoni, D. B. (1998). Recognizing Spoken Words: The Neighborhood Activation Model. *Ear and Hearing*, *19*(1), 1–36. https://doi.org/10.1097%2F00003446-199802000-00001

Malmstrom, J. (1975). Psycholinguistics and Reading. *Elementary English*, *52*(3), 316–319.

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

Meyer, D. E., Schvaneveldt, R. W., & Ruddy, M. G. (1974). Functions of graphemic and phonemic codes in visual word-recognition. *Memory & Cognition*, *2*(2), 309–321. https://doi.org/10.3758/BF03209002

Morais, J. (2021). The phoneme: A conceptual heritage from alphabetic literacy. *Cognition*, *213*, 104740. https://doi.org/10.1016/j.cognition.2021.104740

Morais, J., & Kolinsky, R. (2019). Phonological Abilities in Fully Illiterate Adults. In D. Perin (Ed.), *The Wiley Handbook of Adult Literacy* (1st ed., pp. 41–61). Wiley. https://doi.org/10.1002/9781119261407.ch2

Nielsen, K. (2011). Specificity and abstractness of VOT imitation. *Journal of Phonetics*, *39*(2), 132–142. https://doi.org/10.1016/j.wocn.2010.12.007

Pattamadilok, C., Knierim, I. N., Kawabata Duncan, K. J., & Devlin, J. T. (2010). How does learning to read affect speech perception? *Journal of Neuroscience*, *30*(25), 8435–8444. https://doi.org/10.1523/JNEUROSCI.5791-09.2010

Pattamadilok, C., Morais, J., Ventura, P., & Kolinsky, R. (2007). The locus of the orthographic consistency effect in auditory word recognition: Further evidence from French. *Language and Cognitive Processes*, *22*(5), 700–726. https://doi.org/10.1080/01690960601049628

Pattamadilok, C., Perre, L., Dufau, S., & Ziegler, J. C. (2009). On-line Orthographic Influences on Spoken Language in a Semantic Task. *Journal of Cognitive Neuroscience*, *21*(1), 169–179. https://doi.org/10.1162/jocn.2009.21014

Pattamadilok, C., Perre, L., & Ziegler, J. C. (2011). Beyond rhyme or reason: ERPs reveal task-specific activation of orthography on spoken language. *Brain and Language*, *116*(3), 116–124. https://doi.org/10.1016/j.bandl.2010.12.002

Perre, L., Pattamadilok, C., Montant, M., & Ziegler, J. C. (2009). Orthographic effects in spoken language: On-line activation or phonological restructuring? *Brain Research*, *1275*, 73–80. https://doi.org/10.1016/j.brainres.2009.04.018

Perre, L., & Ziegler, J. C. (2008). On-line activation of orthography in spoken word recognition. *Brain Research*. https://doi.org/10.1016/j.brainres.2007.10.084

Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: The CDP+ model of reading aloud. *Psychological Review*, *114*(2), 273–315. https://doi.org/10.1037/0033-295X.114.2.273

Perry, C., Ziegler, J. C., & Zorzi, M. (2010). Beyond single syllables: Large-scale modeling of reading aloud with the Connectionist Dual Process (CDP++) model. *Cognitive Psychology*, *61*(2), 106–151. https://doi.org/10.1016/j.cogpsych.2010.04.001

Perry, C., Ziegler, J. C., & Zorzi, M. (2013). A Computational and Empirical Investigation of Graphemes in Reading. *Cognitive Science*, *37*(5), 800–828. https://doi.org/10.1111/cogs.12030

Petrova, A., Gaskell, M. G., & Ferrand, L. (2011). Orthographic consistency and word-frequency effects in auditory word recognition: New evidence from lexical decision and rime detection. *Frontiers in Psychology*, *2*(OCT), 1–11. https://doi.org/10.3389/fpsyg.2011.00263

Poeppel, D. (2015). Speech Perception. In *Brain Mapping* (pp. 429–434). Elsevier. https://doi.org/10.1016/B978-0-12-397025-1.00264-5

Pulvermüller, F. (1999). Words in the brain’s language. *Behavioral and Brain Sciences*, *22*(2), 253–279. https://doi.org/10.1017/S0140525X9900182X

Roelofs, A. (2004). Seriality of phonological encoding in naming objects and reading their names. *Memory & Cognition*, *32*(2), 212–222. https://doi.org/10.3758/BF03196853

Roelofs, A. (2006). The influence of spelling on phonological encoding in word reading, object naming, and word generation. *Psychonomic Bulletin & Review*, *13*(1), 33–37. https://doi.org/10.3758/BF03193809

Seidenberg, M. S., & Tanenhaus, M. K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory*, *5*(6), 546–554. https://doi.org/10.1037/0278-7393.5.6.546

Stone, G. O., Vanhoy, M., & Orden, G. C. V. (1997). Perception Is a Two-Way Street: Feedforward and Feedback Phonology in Visual Word Recognition. *Journal of Memory and Language*, *36*(3), 337–359. https://doi.org/10.1006/jmla.1996.2487

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. https://doi.org/10.1037/h0054651

Taft, M. (2011). Orthographic influences when processing spoken pseudowords: Theoretical implications. *Frontiers in Psychology*, *2*(JUN), 1–7. https://doi.org/10.3389/fpsyg.2011.00140

Tunmer, W. E., & Rohl, M. (1991). Phonological Awareness and Reading Acquisition. In D. J. Sawyer & B. J. Fox (Eds.), *Phonological Awareness in Reading: The Evolution of Current Perspectives* (pp. 1–30). Springer. https://doi.org/10.1007/978-1-4612-3010-6_1

Ventura, P., Morais, J., Pattamadilok, C., & Kolinsky, R. (2004). The locus of the orthographic consistency effect in auditory word recognition. *Language and Cognitive Processes*, *19*(1), 57–95. https://doi.org/10.1080/01690960344000134

Walton, P. D., & Walton, L. M. (2002). Beginning Reading by Teaching in Rime Analogy: Effects on Phonological Skills, Letter-Sound Knowledge, Working Memory, and Word-Reading Strategies. *Scientific Studies of Reading*, *6*(1), 79–115. https://doi.org/10.1207/S1532799XSSR0601_04

Ziegler, J. C., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin and Review*, *5*(4), 683–689. https://doi.org/10.3758/BF03208845

Ziegler, J. C., Ferrand, L., & Montant, M. (2004). Visual phonology: The effects of orthographic consistency on different auditory word recognition tasks. *Memory and Cognition*, *32*(5), 732–741. https://doi.org/10.3758/BF03195863

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, *131*(1), 3–29. https://doi.org/10.1037/0033-2909.131.1.3

Chapter 2 – The influence of literacy on auditory language perception

In this chapter, I pose the following research questions:

RQ1 – Do Spanish-speaking and French-speaking adults rely on phoneme-to-grapheme mappings in auditory language perception?

RQ2 – Do Spanish-speaking children rely on phoneme-to-grapheme mappings in auditory language perception?

RQ3 – Can the Orthographic Consistency Effect (OCE) also be observed in a language with an alphabetic writing system with only a few inconsistent phoneme-to-grapheme mappings?

RQ4 – Are there cross-linguistic differences in the extent to which listeners of opaque or transparent languages rely on orthographic information in auditory language perception?

In Experiment 1, I addressed RQ1, RQ3, and RQ4 employing an auditory lexical decision task (LDT) administered to 30 L1 Spanish and 30 L1 French speaking adults. The stimulus lists consisted of words and pseudowords that either contained consistently-spelled phonemes (e.g., /p/ for both Spanish and French) or inconsistently-spelled phonemes (e.g., /f/ for French and /b/ for Spanish) in the first position. I expected that an OCE, as it has also been shown at the phonemic level in auditory language perception. Consequently, I also expected that the OCE would be present in a transparent language like Spanish since the effect occurs in the prelexical stage of auditory processing, and therefore depends on the sub-lexical structure. In other words, what matters is that the individual mappings for those phonemes to graphemes are not one to one, not the overall structure of the language. Concerning cross-linguistic differences, I expected that the OCE would be stronger in French, impacting both word and pseudoword processing to the same extent, given the ubiquitous opaque mappings in the language. As for Spanish, I expected a stronger OCE in word than in pseudoword recognition, as the word would have stronger activation of the phonological representation, whereas pseudowords would simply engage decoding. Overall, the results showed an auditory OCE in French pseudoword processing and Spanish word processing. In Experiment 2, I adopted the same experimental design and tested 45 Spanish L1 seven-year-old schoolers to answer RQ3. The results showed an OCE in both word and pseudoword recognition. Children were, in fact, more accurate at recognizing consistent than inconsistent items.

2.1 – Theoretical motivation for Experiment 1

Previous research on the Orthographic Consistency Effect (OCE) in speech processing showed that words with consistent rimes (e.g., English /əʊb/ as in *gl<u>obe</u>* or *pr<u>obe</u>*) are processed faster and more accurately than words with inconsistent rimes (e.g., English /eɪm/ as in *cl<u>aim</u>* or *fl<u>ame</u>*) in auditory LDTs (Pattamadilok, Morais, et al., 2009; Petrova et al., 2011; Ventura et al., 2004; Ziegler et al., 2008; Ziegler & Ferrand, 1998), to some extent in shadowing tasks (Pattamadilok, De Morais, et al., 2011), and in rime detection tasks (Petrova et al., 2011; Ziegler et al., 2004). The reason why such an effect occurs in auditory language perception is that inconsistent syllables co-activate all their orthographic representations which compete for the final lexical access, consequently delaying the process (Muneaux & Ziegler, 2004).

Perry (2003) argued that inconsistent phoneme-to-grapheme mappings generate a stronger OCE than those at the syllable/rime level since he found a strong OCE (*feedback consistency effect*) in a visual LDT in which the stimuli were manipulated on their phoneme-to-grapheme mappings. His claim was also based on the importance of the phonemic level in writing tasks such as spelling. Similarly, there is evidence of the OCE in studies that addressed different psycholinguistic aspects (e.g., phonological awareness) in the auditory modality in which the manipulation was at the phonemic level. For example, metaphonological tasks like phoneme deletion tasks showed an OCE at the phonemic level: English listeners manipulate individual consistent phonemes more effectively than inconsistent phonemes (Castles et al., 2003). Concretely, it is easier to delete the phoneme /rə/ from *struggle* then /wə/ from *squabble* for L1-English speakers. This is because /rə/ has only one orthographic representation (i.e., <r>) while /wə/ has multiple spellings (i.e., <qu>, <w>, etc.). Additionally, English listeners recognize consistent phonemes more accurately and faster than inconsistent phonemes in the word-initial position (Cutler et al., 2010). However, no study has ever investigated if consistent phoneme-to-grapheme mappings affect auditory language perception.

The first aim of Experiment 1 is, therefore, to understand whether words and pseudowords with only consistent phonemes are processed faster and more accurately than words and pseudowords with inconsistent *phoneme*s, as is the case with stimuli containing consistent or inconsistent *syllables*. In other words, Experiment 1 explores whether the (in)consistency of individual phoneme-to-grapheme mappings influences auditory language perception. The relevance of this question is expressed in models of single-word processing encompassing the phonology-orthography interface, such as the bimodal interactive activation model (Grainger & Ferrand, 1996). This model posits that auditory input activates the phonological sub-lexical units which automatically co-activate the respective orthographic sub-lexical units which contribute to the bottom-up process of lexical access. The facilitatory effect of consistent with respect to inconsistent phonemes should, therefore, result in faster lexical processing while inconsistent phonemes would slow down the process since they co-activate all their orthographic representations. Syllables and rimes are indeed sub-lexical phonological units, but phonemes are the basic linguistic units in phonology. Therefore, understanding whether the interaction between the basic linguistic units in phonology and the basic units in orthography (i.e., graphemes) contributes to speech processing is important from a theoretical point of view.

Another important issue that has not been thoroughly addressed by previous research is whether and to what extent speakers of fairly transparent languages are sensitive to the rare irregularities in their spelling system during auditory lexical processing. So far, Portuguese is the only transparent language for which the OCE has been tested and demonstrated (Ventura et al., 2004). Portuguese, however, has a richer phonological inventory and more inconsistencies in phoneme-to-grapheme mappings than, for example, Spanish (Seymour et al., 2003). This gap in previous literature raises the important theoretical question of whether the orthographic representations of individual phonemes are relevant only for those listeners of languages with many inconsistent phoneme-to-grapheme mappings or whether this is a more common process that occurs regardless of the opacity of the alphabetic language in question. If the OCE underlies a common mechanism deriving from literacy in alphabetic languages, it would then empirically imply that consistent words and pseudowords are processed more rapidly and more accurately than inconsistent auditory stimuli even in a language with a very small number of inconsistent phoneme-to-grapheme mappings. More generally, understanding whether the OCE is detected in a transparent language like Spanish would provide a fuller picture of the impact of literacy on auditory language perception.

Finally, about the last point, I explore possible cross-linguistic differences of the OCE in adults. For instance, Pattamadilok and colleagues (2007) repeated the study of Ventura and colleagues (2004) using French (instead of Portuguese) as the target language. They found that, like in Portuguese, in French the OCE occurred in the auditory LDT but not in shadowing. Yet, the OCE affected both word and pseudoword processing in French, while it only affected word processing in Portuguese (Ventura et al., 2004). Taken together, these diverging results suggest that French listeners may rely more heavily on orthographic information in auditory language perception than Portuguese listeners. Pattamadilok and colleagues argue that French listeners rely more on orthographic representations in auditory language perception because they help with the selection of the correct lexical entry (in French, /pɛ̃/ can refer to *pain* ‘bread’ or *pin* ‘pine’). Alternatively, the authors suggest that the link between phonology and orthography might be stronger in French than Portuguese because of the different orthographic transparency in the two languages. This hypothesis is based on the reading acquisition literature that shows that children who learn to read in languages with an opaque writing system establish connections not only between individual phonemes to the correspondent graphemes but also between sounds and spelling of more complex linguistic units (Seymour et al., 2003). However, studies on reading acquisition like Seymour and colleagues found similar patterns in reading acquisition between French- and Portuguese-speaking children. A cross-linguistic comparison between French and Spanish, therefore, is more appropriate to understand whether the sensitivity to the (in)consistencies in phoneme-to-grapheme mappings is higher in listeners of an opaque language like French compared to listeners of a transparent language like Spanish in auditory language perception, as it was found in reading acquisition.

For Experiment 1, the following hypotheses were formulated:

(a) The OCE is rooted at the phonemic level. A consistent word or pseudoword should be recognized faster than an inconsistent one given that each phoneme would co-activate only one orthographic representation. Conversely, inconsistent words but also pseudowords should be cognitively costlier to process (i.e., longer response times) because any inconsistent sound will co-activate multiple competing orthographic representations. Therefore, I expected that consistent phoneme-to-grapheme mappings facilitate auditory language perception.

(b) I hypothesize that even in highly transparent languages, inconsistent phonemes elicit processing costs (i.e., longer RTs) because they co-activate multiple orthographic representations that compete and, consequently, slow down lexical access. Thus, the OCE should occur in Spanish even though it presents only a few inconsistent phoneme-to-grapheme mappings.

(c) From a cross-linguistic perspective, the OCE should be stronger in French because it has more inconsistent phoneme-to-grapheme mappings than Spanish which make French listeners more sensitive to orthographic inconsistencies in general (Pattamadilok et al., 2007). Both French word and pseudowords processing should show a comparable OCE. Conversely, Spanish should show a comparable pattern to what Portuguese showed in Ventura and colleagues (2004): the OCE should be weaker especially in Spanish pseudoword processing because the competing orthographic representations related to inconsistent phonemes do not participate in any lexical selection.

2.2 – Methodology

2.2.1 – Participants

Thirty native French speakers (17 female, M_{age}= 21.83 years, *SD*=2.13) and 30 native Spanish speakers (17 female, M_{age}= 24.23 years, *SD*=3.11) participated in Experiment 1. All participants had university-level education; their ages ranged from 18 to 30 years.

Based on self-report, participants of both groups had been exposed to their native language since birth and they all predominantly used their native language in everyday life. No specific learning impairments and no hearing or uncorrected vision problems were reported. French and Spanish participants were also matched on non-verbal IQ (p = 0.371), as measured by the Kaufman Brief Intelligence Test (Kaufman & Kaufman, 2004). The French cohort was recruited from the University of Bordeaux (France). Native Spanish speakers were recruited from the participant pool of the BCBL (Spain). Participants received monetary compensation for their participation and all signed consent forms, previously approved by the BCBL's Ethics Committee and by the Bioethics Commission of the University of Barcelona, before starting the experiment.

2.2.2 – Materials

Stimuli were recorded by male native speakers of French and Spanish in a sound-attenuating chamber using a Marantz PMD 671 digital recorder with a Sennheiser ME65 microphone. All stimuli were scaled to 65dB and a 50ms interval of silence was added to the beginning of each audio file to allow for sufficient loading time in the experimental software. The stimulus list for each language consisted of 60 words and 60 pseudowords. For each lexicality condition (words; pseudowords), 50% of the stimuli contained phonemes with only one possible spelling, that is consistent. In the other 50% of the stimuli, at least the first phoneme in each item could be spelled in more than one way, which is inconsistent. Inconsistent French items contained up to five inconsistent phonemes, whereas Spanish items contained up to three. The discrepancy between the two languages was due to differences in the opacity of French and Spanish: French words commonly include several inconsistent phonemes; Spanish is quite transparent and has fewer words with multiple inconsistent phonemes. The stimulus lists needed to contain words that were also known by children because the intention was to employ the same stimuli for both Experiment 1 and Experiment 2.

All inconsistent target phonemes were consonants. Pseudowords were created from the word lists by changing the consonants at the beginning of each syllable, respecting phonotactic constraints (see Table 2).

Table 2. Example of stimuli for French and Spanish.

	French		Spanish	
	Consistent	Inconsistent	Consistent	Inconsistent
Words	porte /pɔʁt/	flaque /flak/	fruta /fruta/	brazo /braθo/
Pseudowords	/lɔʁv/	/slas/	/lursa/	/kraxo/

Note: Inconsistent phonemes highlighted in bold

The Spanish and French stimuli were matched on the same variables (see Appendix A). Cross-linguistically, words were matched on frequency (Zipf's log frequency), number of letters, and the duration of the audio recordings (see Appendix B). These variables were retrieved from the

Lexique database (New et al., 2004) for French and from the EsPal database (Duchon et al., 2013a) for Spanish. The phonological and orthographic neighborhood densities of words and pseudowords of both languages were determined with CLEARPOND (Marian et al., 2012b). Consistent and inconsistent pseudowords were matched on these variables within languages. Across languages, pseudowords were matched on the mean phonological neighborhood (see Appendix B).

2.2.3 - Apparatus and procedure

The experiment was administered on a laptop computer (HP EliteBook Folio 1040 G3) using OpenSesame software (version 3.2.4; Mathôt et al., 2012). Auditory stimuli were presented over headphones (Sennheiser GSP 350). The experiment was run in sound-attenuating chambers at the University of Bordeaux and at the BCBL.

Participants were tested on an auditory LDT, in which they had to respond via key press whether what they heard was a real word or not in their L1. Participants were instructed to respond as quickly and as accurately as possible. The respective keys were labeled on the keyboard; half of the participants pressed the left key for words and the other half pressed the right key for words. Before starting the main part of the experiment, the participants completed a practice phase with 10 extra items (5 words, 5 pseudowords). The participants only received feedback on their responses in the practice phase. During both the practice phase and the main experiment, each trial started with a fixation point appearing for 500ms, which was followed by the auditory stimulus. RT measurement started at stimulus onset and ended with the key press. After participants gave their responses, the next trial was automatically initiated. The items were presented in randomized order across participants. The task lasted approximately 15 minutes.

2.3 – Results

Due to technical problems, two Spanish participants did not fully complete the task, such that 0.53% of the entire dataset was missing. Both language groups performed at ceiling on accuracy so no meaningful statistical analysis could be performed. Consequently, only RT data were analyzed, and the analysis included only correct responses (95.88% of the entire dataset).

Two consistent French pseudowords—/myv/ and /sadʁ/—were excluded from further analyses because their accuracy rates were below chance (1.68% of the entire dataset). The remaining consistent and inconsistent pseudowords remained matched on all confounding variables. After excluding these data points, excessively long (> 3000ms) and short (< 150ms) RTs were removed (0.12% of the correct responses). Next, RTs above and below 3SD from the mean were removed on a by-participant basis (5.35% of correct responses). Table 3 reports descriptive statistics for RTs.

Table 3. Mean RT of the LDT for both languages by lexicality and consistency (RT in ms).

	French		Spanish	
	Word	Pseudoword	Word	Pseudoword
Consistent	973 (201)	1086 (213)	957 (159)	1067 (222)
Inconsistent	990 (182)	1133 (213)	998 (179)	1063 (202)
Δ	17	47	41	-4

Note: Delta expresses the difference in milliseconds between inconsistent and consistent items. SD in parentheses.

The main analysis was run in RStudio (version 1.3.1073; RSTudio Team, 2020) using the *lme4* package (Bates et al., 2015). RT data were analyzed using a generalized linear mixed-effects model (GLMM) with the *glmer* function of *lme4*. The choice of a GLMM over a linear mixed-effects model (LMM) or ANOVA is because raw RTs cannot satisfy the assumption of normal distribution which is required by linear regressions. The distribution problem could have been tackled with a non-linear transformation (e.g., logarithmic) of the raw RT. Such transformations, however, can lead to misinterpretations of the results (Lo & Andrews, 2015). The GLMM was run based on the assumption of a Gamma distribution of the data with an identity link, which is in line with previously-proposed best practice. The predicted variable was *RT* (expressed in ms) and the predictors were *Consistency* (consistent = -1, inconsistent = 1), *Lexicality* (word = -1, pseudoword = 1), and *Language* (French = -1, Spanish = 1) which were contrast coded in line with the best practice in the usage of (G)LMM (Schad et al, 2020). The three predictors were linked with a three-way interaction term including lower-level interactions. Random intercepts for *Participant* and *Stimulus* with by-*Participant* random slopes for *Lexicality* and *Consistency* were also included. The optimizer BOBYQA was applied to the model to solve convergence problems (Powell, 2009).

The model detected significant effects for *Consistency*, *Lexicality*, *Language*, a significant interaction between *Lexicality* and *Language,* and a significant three-way interaction between *Consistency*, *Lexicality,* and *Language* as shown in Table 4.

Table 4. Main GLMM output from the RT analysis.

Fixed effects		β	SE	*t*	*Pr (>|t|)*
Intercept		1078.98	4.14	260.87	<.001
Consistency		12.94	3.86	3.35	.0008
Lexicality		62.54	4.15	15.07	<.001
Language		-16.38	3.76	-4.34	<.001
Consistency*Lexicality		-2.49	3.93	-0.63	.527
Consistency*Language		-3.76	3.90	-0.96	.336
Lexicality*Language		-13.1	4.97	-2.63	.0085
Consistency*Lexicality*Language		-9.18	4.41	-2.08	.0372
Random effects	*Group*	*Variance*	*SD*	*Correlation*	
Item	Intercept	2.583e+03	50.83		
Participant	Intercept	2.292e+03	47.87		
	Lexicality	5.369e+02	23.17	0.36	
	Consistency	1.334e+02	11.55	0.11	0.07

Note: SE= standard error; Pr (>|t|) p-value calculated from the t-value with lmerTest package (Kuznetsova et al., 2017)*; SD = standard deviation.*

The three-way interaction between *Consistency*, *Lexicality,* and *Language* indicated that the OCE emerged in both French and Spanish but to a different extent. As visualized in Figure 4, French participants recognized both consistent words and consistent pseudowords faster than inconsistent words and inconsistent pseudowords respectively; Spanish participants recognized consistent words faster than inconsistent words, but no trend for a consistency effect emerged in pseudoword recognition.

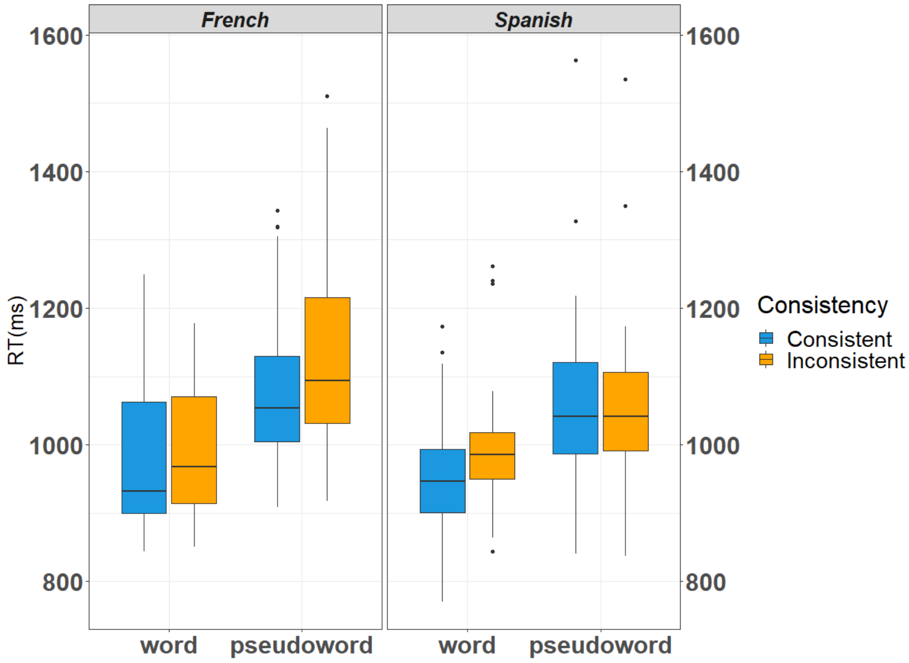


Figure 4. Distribution of the RTs (grouped by participants) across conditions and languages.

The three-way interaction was further investigated through multiple comparisons using the *ghlt* function of the *multcomp* R package (Hothorn et al., 2008) and the e*mm* function of the *emmeans* R package (Lenth, 2021). More specifically, the pairwise contrasts were between *Consistency* and the interaction term of *Lexicality* and *Language*. The alpha level was adjusted with Holm correction. Table 5 summarizes the output of the post-hoc analysis.

Table 5. Post-hoc analysis by multiple comparisons.

French	β	SE	*z*-value	*Pr(>|z|)*
Word	-19.99	16.11	-1.24	.215
Pseudoword	-46.78	15.82	-2.96	.0031

Spanish	β	SE	*z*-value	*Pr(>|z|)*
Word	-41.70	16.43	-2.54	.0111
Pseudoword	4.98	16.16	0.31	.758

Note: SE = standard error; Pr (>|t|) *p*-value calculated from the *t*-value with the lmerTest package (Kuznetsova et al., 2017); SD = standard deviation.

Overall, multiple comparisons show that the OCE affects pseudoword processing in French and word processing in Spanish. This pattern is reflected in the mean RTs across conditions: the biggest differences in mean RTs within the *Consistency* condition are in French pseudowords (Δ = 47ms) and Spanish words (Δ = 41ms). The mean RT difference between French inconsistent words and consistent words (Δ = 17ms) only reflects a numerical tendency for the OCE. As for Spanish pseudowords, the mean RT difference between inconsistent and consistent pseudowords (Δ = -4ms) did not show any numerical pattern for the OCE.

2.4 – Summary of Experiment 1

Experiment 1 investigated whether orthographic (in)consistencies at the phonemic level affect auditory language processing in adults. For the first time, I also examined whether the OCE is present in Spanish as a representative of orthographically transparent languages compared to French as representative of orthographically opaque languages. Finally, it also addressed cross-linguistic differences.

First, French and Spanish listeners identified words and pseudowords at ceiling irrespective of their (in)consistency. Even though accuracy is often considered an important variable in the study of the OCE, it did not prove to be a sensitive measure for the current data. One reason for ceiling performance could be the high frequency of the stimuli (mean French word frequency: 59.62/million; mean Spanish word frequency: 54.73/million). Selecting such high-frequency stimuli was necessary for Experiment 2, in which children complete the same LDT using the same materials. Conversely, reaction time analyses showed that inconsistencies at the individual phonemic level impacted language processing. This novel finding extends what previous research investigated regarding the OCE at the suprasegmental level (Pattamadilok et al., 2009; Ventura et al., 2004; Ziegler et al., 2004). Experiment 1 provides, therefore, the first empirical evidence that the OCE affects word processing in Spanish, a language with very few inconsistent phoneme-to-grapheme mappings. Consequently, orthographic inconsistencies affect word processing regardless of the degree of opacity of the language—even in languages with a negligible number of inconsistencies.

Second, the systematic comparison of the OCE in French and Spanish revealed important differences in the way orthographic inconsistencies at the phonemic level affect auditory language perception in opaque and transparent languages. In French, only a numerical tendency shows that listeners recognized inconsistent words more slowly than consistent words. This statistically non-significant difference in the recognition speed between consistently-spelled and inconsistently-spelled French words (Δ = 17ms) is numerically smaller than the difference found in previous studies manipulating consistency of the rime level (e.g., Pattamadilok et al., 2007 with Δ = 61ms; Ziegler & Ferrand, 1998 with Δ = 62ms). A possible explanation for the non-significance of the OCE in French word recognition is that French listeners are more sensitive to orthographic inconsistencies at the rime level compared to orthographic inconsistencies at the phonemic level. In other words, it seems that the sound-to-spelling mappings of bigger units (i.e., rimes) are more relevant than sound-to-spelling mappings of individual phonemes for languages with opaque writing systems during auditory word recognition (see Grain-Size Theory in Chapter 1). However, this is not necessarily the case for pseudoword recognition. Finding an OCE in French pseudoword recognition is in line with what Pattamadilok and colleagues (2007) found. Yet, it is in contrast with other studies on the OCE in French (e.g., Ziegler et al., 2004; Ziegler & Ferrand, 1998). Numerically, the difference in the response latencies between consistent and inconsistent French pseudowords (Δ = 47ms) is larger here than in Pattamadilok and colleagues (2007; Δ = 35). A possible explanation for this numerical difference could be that inconsistent phoneme-to-grapheme mappings are more relevant than those of bigger linguistic units during phonological decoding because phonological decoding in general is based on phonemes and not on syllables or rimes (Perry, 2003). However, this hypothesis should be tested in a dedicated study in which inconsistency is tested either at the phonemic or at the rime level.

Spanish listeners showed the opposite pattern as French listeners since orthographic consistencies impacted auditory word recognition but not auditory pseudoword recognition. Thus, the results for Spanish pseudowords appear at odds with the explanation I propose for auditory pseudoword processing in French and with the consequent pre-lexical origin of the OCE. Yet, listeners of languages with more transparent orthographies—such as Spanish—may rely less on orthography than listeners of languages with opaque orthographies—such as French (Pattamadilok et al., 2007). Indeed, the present results are in line with those of Ventura and colleagues (2004) who found an OCE only in auditory word processing and not pseudoword processing in Portuguese, a language which is closer to Spanish than to French on the orthographic depth continuum. Alternatively, it is possible that I did not find a strong OCE in Spanish pseudoword processing because, as suggested by Taft (2011), it probably does not matter what competing orthographic representations are activated (e.g., whether a pseudoword with /b/ activates <b> or <v>), particularly if there is a more dominant spelling for a given inconsistent phoneme, whenever there is not any lexical access.

In conclusion, Experiment 1 suggests that the role of orthography in auditory language perception varies by language. One aspect that is hereafter covered in this thesis is the role of phoneme-to-grapheme mappings in young readers' auditory language perception (Experiment 2). Only children who are learning to read can provide an insight into how phoneme-to-grapheme connections are established, following the co-activation account, or phoneme representations are restructured after reading acquisition, following the phonological restructuring account, as outlined in the following section.

2.5 – Theoretical motivation for Experiment 2

Learning to read and write in a language with an alphabetic writing system means establishing connections between sounds and letters (i.e., graphemes). The process has been named *phonological recoding* (Ziegler & Goswami, 2005) and it is characterized by different steps in which children apply statistical learning strategies in phoneme-to-grapheme mapping determination (Treiman & Kessler, 2013). Phoneme-to-grapheme mappings take a unique role at this developmental stage but have never been systematically manipulated in speech processing studies conducted with children. The main aim of Experiment 2, therefore, was to determine whether these mappings play a role in children's speech processing. Like in Ventura and colleagues (2007), I expected that children at an early stage of reading acquisition would strongly rely on phoneme-to-grapheme mappings during speech perception. Differently from previous studies, the selection of Spanish as a target language with very few inconsistent phoneme-to-grapheme mappings further contributes to the field by showing whether the patterns found in previous research also apply in a transparent language. Previous research was based primarily on French or English (i.e., languages with opaque alphabetic writing systems) and Portuguese, a language with quite rich a phonemic inventory and with more inconsistencies in phoneme-to-grapheme mappings than Spanish. From a behavioral perspective, the difference in orthographic transparency between Portuguese and Spanish becomes evident when comparing reading acquisition in the two languages: Spanish-speaking children (from grade 1 [age 6 years] to 4 [age 10 years]) make fewer phonological errors during reading than grade-matched Portuguese-speaking children (Defior et al., 2002).

The initially-planned second goal of Experiment 2 was to systematically compare the magnitude of the OCE at the phonemic level between Spanish, as a representative of transparent languages, and French, a representative of opaque languages in young readers. Previous cross-linguistic developmental studies investigated the impact of literacy acquisition on different linguistic aspects such as phonemic awareness, reading skills or vocabulary growth (Duncan et al., 2013; Metsala, 1999). To date, Ventura and colleagues (2007) and Pattamadilok and colleagues (2009) were the only ones who investigated the developmental trajectory of the OCE in speech processing, the first for Portuguese and the second for French children. Since both studies reported the same pattern in both populations, the systematic cross-linguistic comparison in Experiment 2 aimed to understand whether this pattern is common to learns of alphabetic languages in a more extended sense. Unfortunately, due to the COVID-19 pandemic it was not possible to address this last issue because of many restrictions related to French schools.

2.6 – Methodology

2.6.1 – Participants

Forty-five Spanish second graders (female = 19, M_{age} = 7.6, range = 7;3-8;3) from a school in Vitoria-Gasteiz (Spain) participated in the study under their caregivers' permission. Besides Spanish, children were exposed to Basque and English at school. As confirmed by parental report, Spanish—the children's L1—was their dominant language. Teachers reported no learning or reading difficulties for the participants. By the time of the testing (April 2019), the children had received one year and a half of reading instruction. The study was approved by the BCBL's Ethics Committee and by the Bioethics Commission of the University of Barcelona.

2.6.2 – Materials

The stimulus list was the same as the Spanish one used in Experiment 1. According to the EsPal Database (Duchon et al., 2013a), words had an average Age of Acquisition of 5.14 years, and this value was matched across the consistency conditions (*p* = 0.47), ensuring that children would know the words in the stimulus list.

2.6.3 – Apparatus and procedure

The apparatus and the procedure were the same as in Experiment 1. To avoid fatigue, there was a break every 30 trials. The children were tested in a silent room at school. The task lasted approximately 20 minutes.

2.7 – Results

Due to a technical problem, 3.56% of the data were lost during the experiment. The overall accuracy rate of the participants was 80.31%. Two participants performed at chance level and they were therefore excluded from the analysis. Three words (i.e., *doña*, *potro* and *kayak*) were also excluded from the analysis because of an accuracy rate below 50% (2.48% of the entire dataset). The stimuli remained matched on all confounding variables within the *Consistency* condition. Two measures were taken into account: RT and Accuracy.

Regarding RTs, the analysis was run only on correct responses (85.89% of the entire dataset). The dataset was trimmed by removing excessively long (>5000ms) and short trials (<150ms), which represent 0.85% of the correct responses. On a by-participant basis, trials above and below 3SD of the mean RT were also excluded from the analysis (1.74% of the correct responses). Table 6 summarizes the mean RTs across conditions.

Table 6. Mean RT of the LDT by lexicality and consistency (RT in ms).

	Word	Pseudoword
Consistent	1531 (480)	1804 (557)
Inconsistent	1541 (482)	1785 (586)
Δ	10	-47

Note: Delta expresses the difference in ms between inconsistent and consistent items. SD in parentheses

The main analysis was run in RStudio (version 1.3.1073; RSTudio Team, 2020) using the *lme4* package (Bates et al., 2015). RT data was analyzed using a generalized linear mixed-effects model (GLMM) with the *glmer* function of *lme4*. The predicted variable was *RT* (expressed in ms) and the predictors were *Consistency* (consistent = -1, inconsistent = 1), *Lexicality* (word = -1, pseudoword = 1), and *Language* (French = -1, Spanish = 1) which were contrast coded. The three predictors were linked with a three-way interaction term including lower-level interactions. Random intercepts for *Participant* and *Stimulus* with by-*Participant* random slopes for *Lexicality* and *Consistency* were also included. The optimizer BOBYQA was applied to the model to solve convergence problems (Powell, 2009).

The analysis shows only a significant effect of *Lexicality* (β = 132.33, *t* = 13.521, *p* < 0.001) whereas neither a significant effect for *Consistency* nor an interaction between *Consistency* and *Lexicality* resulted from the model.

Regarding the analysis on accuracy, only the three words with low accuracy rate were excluded. The logistic mixed-effects model had *Accuracy* (correct response = 1, incorrect response = 0) as the dependent variable with fixed effects for *Consistency* (consistent=-1, inconsistent = 1) and *Lexicality* (word = -1, pseudoword = 1) with an interaction term. The model also included random intercepts for *Subjects* and *Items*, as well as by-subject random slopes for *Consistency* and *Lexicality*. Also in this case, the optimizer BOBYQA was applied to the model to solve convergence problems.

The analysis shows a main effect for *Consistency* (β = -0.27, *z* = -2.38, *p* = 0.0172) with no other effect or interaction, as shown in Table 7.

Table 7: Main logistic mixed-effects model output from the accuracy analysis

Fixed effects		β	SE	*z*	*Pr (>|z|)*
Intercept		2.95	0.14	260.87	<.001
Consistency		-0.27	0.11	-2.38	.0172
Lexicality		-0.13	0.14	-0.88	.377
Consistency*Lexicality		0.20	0.11	1.79	.0733
Random effects	*Group*	*Variance*	*SD*	*Correlation*	
Item	Intercept	0.93	0.96		
Participant	Intercept	0.21	0.46		
	Consistency	0.01	0.11	-0.69	
	Lexicality	0.31	0.56	0.95	-0.42

Note: SE= standard error; Pr (>|z|) p-value calculated from the z-value with the lmerTest package (Kuznetsova et al., 2017)*; SD = standard deviation.*

As shown in Figure 5, children were overall more accurate when recognizing consistent than inconsistent items.

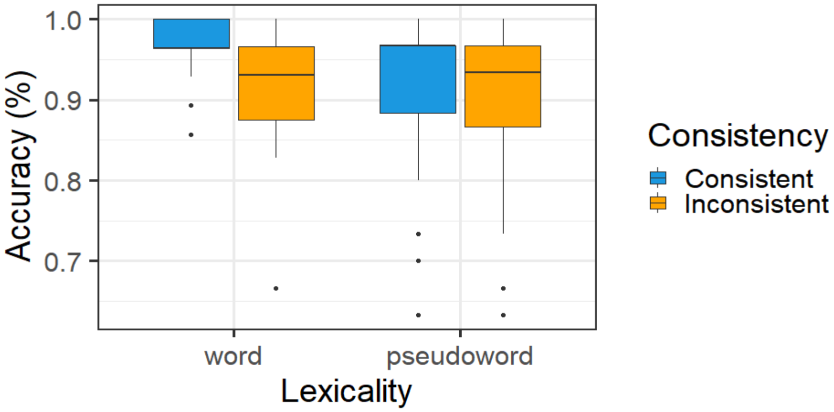


Figure 5. Distribution of accuracy (grouped by participants) across conditions.

2.8 – Summary of Experiment 2

In Experiment 2, I investigated whether Spanish-speaking L1 seven-year-olds at an early stage of reading acquisition show an OCE during language perception. As in Experiment 1, I administered an auditory LDT, in this case to children, in which they had to determine whether the auditorily-presented stimuli were real Spanish words or not. Overall, the RT analysis shows neither a significant difference in the *Consistency* condition nor an interaction between *Consistency* and *Lexicality*. Even though previous studies (e.g., Ventura et al. 2007) reported that second graders were quicker at recognizing both consistent words and pseudowords in this paradigm, I argue that RTs might not always be a reliable measure with young children. In fact, this type of yes/no task has the limitation that it brings more variability in the data than a go/no-go LDT (Moret-Tatay & Perea, 2011). Nevertheless, I opted for this task because I had concrete hypotheses for pseudowords which could not have been tested otherwise.

I found that children were more accurate when recognizing consistent stimuli (i.e., both words and pseudowords) than inconsistent stimuli. This finding is in line with previous research which shows orthography pervasively impacts children's language skills, including language perception. As discussed by Ventura and colleagues (2007), children rely on phonological decoding during reading acquisition to build and strengthen phoneme-to-grapheme mappings. In this stage, children make connections between the already established phonological representations with novel graphemic representations. The present study provides clear evidence of phonological decoding in auditory language perception since the stimulus manipulation was at the phonemic and not at the rime level as in previous research.

2.9 – General discussion

The Orthographic Consistency Effect (OCE)—a processing advantage for words with phonological units presenting unambiguous spellings—has previously been observed in words with consistent compared to inconsistent rime-spellings in orthographically opaque languages. In contrast to previous research, this study manipulated inconsistency at the phonemic rather than the suprasegmental (i.e., rime) level.

The findings of the present study are in line with research on metalinguistic skills showing that listeners more easily manipulate consistent than inconsistent phonemes (Castles et al., 2003) and that they recognize consistent phonemes faster and more accurately than inconsistent phonemes (Cutler et al., 2010). Most importantly, the present findings are in line with the OCE found in the visual modality when the manipulation was also at the phonemic level (Perry, 2003). The phonemic level plays an important role in language processing in general for alphabetically literate people. In fact, phonemic awareness has been shown to be a better predictor of reading acquisition than onset/rime awareness (Hulme et al., 2002). Since literacy mainly means establishing a relationship between individual phonemes and their orthographic representations. The present findings confirm the influence of literacy on auditory language perception.

The present results for Spanish thus suggest that orthographic information is at play during auditory word recognition. When processed auditorily, Spanish words with an inconsistent phoneme co-activate all the orthographic representations of that inconsistent phoneme, slowing down lexical processing. It should be noted that all Spanish inconsistencies, apart from the phoneme /b/, are regulated by orthographic rules. For example, in standard Castilian Spanish, the phoneme /θ/ can be spelled with <c> when it precedes <i> and <e> (e.g., *cita* or *cera*), but corresponds to <z> when it precedes <a>, <o> or <u> (e.g., *zapato*, *zorro* or *zulo*). Thus, this indicates that the competition of the different orthographic representations of an inconsistent phoneme occurs pre-lexically. There would not be any competing orthographic representations for a word like /kasa/ (Spanish for 'home') if the level at which speech processing takes place would not start from the syllabic level because the syllable /ka-/ in Spanish can only be spelled with <ca->[3].

In summary, the findings of Experiment 1 seem to fit the bimodal interactive activation model proposed by Grainger and Ferrand (1996). Inconsistent phonemes co-activate their competing orthographic representations at the sub-lexical level which in turn slow down lexical access. However, the present findings show that this co-activation is not the same across languages and/or across lexical conditions (i.e., words or pseudowords). On the one hand, it seems that the phonemic level plays a role in French only when phonological decoding is involved. In fact, the numerical tendency of the OCE in French word processing could suggest that orthographic representations of individual phonemes are less activated because French listeners are probably relying more on larger linguistic units. Yet, when it comes to recognizing a French pseudoword, due to decoding, the orthographic representations of phonemes are more activated and collaborate in the process. On the other hand, the phonemic level seems to play a role in auditory word recognition in Spanish because, contrarily to French, Spanish orthographic inconsistencies exist only at the phonemic level. This pattern of the findings suggests that maybe there is also a top-down process according to which lexical representations influence sub-lexical linguistic representations. Indeed, there is evidence that top-down processes occur during speech processing so that lemmas impact auditory lexical perception (Getz & Toscano, 2019). This top-down process is also supported by the bimodal interactive activation model since, like all interaction models, it posits that there is a joint contribution of the different linguistic levels during auditory language processing. In Spanish-speaking seven-year-olds, however, this top-down process seemed to be less impactful than in Spanish-speaking adults. I argue that this is related to children's limited vocabulary. This means that there are fewer lexical representations that influence the sub-lexical processing of an auditory stimulus. This argument also further supports the claim that the OCE found in children's auditory language

---

[3] Another possible orthographic representation of /ka-/ could be <ka-> in Spanish loan words like "karate".

perception is probably related to an ample usage of phonological decoding. Since there is less lexical contribution, children resort to phonological decoding to recognize a (pseudo)word.

An alternative account to the phonology-orthography co-activation with which the present data can be interpreted is the phonological restructuring account (see e.g., Taft & Hambly, 1985). This account posits that phonological representations are 'imbued' with orthographic information. The restructuring of phonological representations happens during reading acquisition in such a way that consistent phonemes are more 'specified' and inconsistent phonemes are less 'specified' (Goswami, 2000). In the context of the present study, the activation of less specified phonological representation of inconsistent phonemes delayed lexical access in Spanish considering the OCE in word recognition. As for French, there was not a significant delay in lexical access because the phonological representation of individual phonemes in that language might be more specified than those of syllables and rimes. Yet, as already mentioned, this difference of linguistic levels should be addressed in future research.

In the following chapter, I will address the issue of which of the two accounts, namely the co-activation or the phonological restructuring, better explains the impact of literacy on auditory language perception.

References

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Castles, A., Holmes, V. M., Neath, J., & Kinoshita, S. (2003). How does orthographic knowledge influence performance on phonological awareness tasks? *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *56 A*(3), 445–467. https://doi.org/10.1080/02724980244000486

Cutler, A., Treiman, R., & van Ooijen, B. (2010). Strategic Deployment of Orthographic Knowledge in Phoneme Detection. *Language and Speech*, *53*(3), 307–320. https://doi.org/10.1177/0023830910371445

Defior, S., Martos, F., & Cary, L. (2002). Differences in reading acquisition development in two shallow orthographies: Portuguese and Spanish. *Applied Psycholinguistics*, *23*(1), 135–148. https://doi.org/10.1017/s0142716402000073

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

Duncan, L. G., Castro, S. L., Defior, S., Seymour, P. H. K., Baillie, S., Leybaert, J., Mousty, P., Genard, N., Sarris, M., Porpodas, C. D., Lund, R., Sigurðsson, B., Þráinsdóttir, A. S., Sucena, A., & Serrano, F. (2013). Phonological development in relation to native language and literacy: Variations on a theme in six alphabetic orthographies. *Cognition*, *127*(3), 398–419. https://doi.org/10.1016/j.cognition.2013.02.009

Getz, L. M., & Toscano, J. C. (2019). Electrophysiological Evidence for Top-Down Lexical Influences on Early Speech Perception. *Psychological Science*, *30*(6), 830–841. https://doi.org/10.1177/0956797619841813

Goswami, U. (2000). Phonological representations, reading development and dyslexia: Towards a cross-linguistic theoretical framework. *Dyslexia*, *6*(2), 133–151. https://doi.org/10.1002/(SICI)1099-0909(200004/06)6:2<133::AID-DYS160>3.0.CO;2-A

Grainger, J., & Ferrand, L. (1996). Masked orthographic and phonological priming in visual word recognition and naming: Cross-task comparisons. *Journal of Memory and Language*, *35*(5), 623–647. https://doi.org/10.1006/jmla.1996.0033

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous Inference in General Parametric Models. *Biometrical Journal*, *50*(3), 346–363. https://doi.org/10.1002/bimj.200810425

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Brief Intelligence Test* (2nd Editio). AGS, American Guidance Service Circle Pines, MN.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lenth, R. V. (2021). *emmeans: Estimated Marginal Means, aka Least-Squares Means* (1.6.3.) [Computer software]. https://CRAN.R-project.org/package=emmeans

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology*, *6*, 1171. https://doi.org/10.3389/fpsyg.2015.01171

Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). CLEARPOND: Cross-Linguistic Easy-Access Resource for Phonological and Orthographic Neighborhood Densities. *PLoS ONE*, *7*(8), e43230. https://doi.org/10.1371/journal.pone.0043230

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–324. https://doi.org/10.3758/s13428-011-0168-7

Metsala, J. L. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Educational Psychology*, *91*(1), 3–19. https://doi.org/10.1037/0022-0663.91.1.3

Moret-Tatay, C., & Perea, M. (2011). Is the go/no-go lexical decision task preferable to the yes/no task with developing readers? *Journal of Experimental Child Psychology*, *110*(1), 125–132. https://doi.org/10.1016/j.jecp.2011.04.005

Muneaux, M., & Ziegler, J. C. (2004). Locus of orthographic effects in spoken word recognition: Novel insights from the neighbour generation task. *Language and Cognitive Processes*, *19*(5), 641–660. https://doi.org/10.1080/01690960444000052

New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. In *Behavior Research Methods, Instruments, and Computers* (Vol. 36, Issue 3, pp. 516–524). Psychonomic Society Inc. https://doi.org/10.3758/BF03195598

Pattamadilok, C., De Morais, J. J., & Kolinsky, R. (2011). Naming in noise: The contribution of orthographic knowledge to speech repetition. *Frontiers in Psychology*, *2*(DEC), 1–12. https://doi.org/10.3389/fpsyg.2011.00361

Pattamadilok, C., Morais, J., De Vylder, O., Ventura, P., & Kolinsky, R. (2009). The orthographic consistency effect in the recognition of French spoken words: An early developmental shift from sublexical to lexical orthographic activation. *Applied Psycholinguistics*, *30*(3), 441–462. https://doi.org/10.1017/S0142716409090225

Pattamadilok, C., Morais, J., Ventura, P., & Kolinsky, R. (2007a). The locus of the orthographic consistency effect in auditory word recognition: Further evidence from French. *Language and Cognitive Processes*, *22*(5), 700–726. https://doi.org/10.1080/01690960601049628

Pattamadilok, C., Morais, J., Ventura, P., & Kolinsky, R. (2007b). The locus of the orthographic consistency effect in auditory word recognition: Further evidence from French. *Language and Cognitive Processes*, *22*(5), 700–726. https://doi.org/10.1080/01690960601049628

Perry, C. (2003). A phoneme-grapheme feedback consistency effect. *Psychonomic Bulletin and Review*, *10*(2), 392–397. https://doi.org/10.3758/BF03196497

Petrova, A., Gaskell, M. G., & Ferrand, L. (2011). Orthographic consistency and word-frequency effects in auditory word recognition: New evidence from lexical decision and rime detection. *Frontiers in Psychology*, *2*(OCT), 1–11. https://doi.org/10.3389/fpsyg.2011.00263

Powell, M. J. D. (2009). The BOBYQA algorithm for bound constrained optimization without derivatives. *Department of Applied Mathematics and Theoretical Physics, University of Cambridge (UK). Technical Report*, *NA2009/06*.

RSTudio Team. (2020). *RStudio: Integrated Development for R* (1.3.1073) [Computer software]. RStudio, Inc. https://rstudio.com

Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, *110*, 104038. https://doi.org/10.1016/j.jml.2019.104038

Seymour, P. H. K., Aro, M., Erskine, J. M., & Network, collaboration with C. A. A. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, *94*(2), 143–174. https://doi.org/10.1348/000712603321661859

Taft, M., & Hambly, G. (1985). The influence of orthography on phonological representations in the lexicon. *Journal of Memory and Language*, *24*(3), 320–335. https://doi.org/10.1016/0749-596X(85)90031-2

Treiman, R., & Kessler, B. (2013). Learning to Use an Alphabetic Writing System. *Language Learning and Development*, *9*(4), 317–330. https://doi.org/10.1080/15475441.2013.812016

Ventura, P., Morais, J., & Kolinsky, R. (2007). The development of the orthographic consistency effect in speech recognition: From sublexical to lexical involvement. *Cognition*, *105*(3), 547–576. https://doi.org/10.1016/j.cognition.2006.12.005

Ventura, P., Morais, J., Pattamadilok, C., & Kolinsky, R. (2004). The locus of the orthographic consistency effect in auditory word recognition. *Language and Cognitive Processes*, *19*(1), 57–95. https://doi.org/10.1080/01690960344000134

Ziegler, J. C., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin and Review*, *5*(4), 683–689. https://doi.org/10.3758/BF03208845

Ziegler, J. C., Ferrand, L., & Montant, M. (2004). Visual phonology: The effects of orthographic consistency on different auditory word recognition tasks. *Memory and Cognition*, *32*(5), 732–741. https://doi.org/10.3758/BF03195863

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, *131*(1), 3–29. https://doi.org/10.1037/0033-2909.131.1.3

Ziegler, J. C., Petrova, A., & Ferrand, L. (2008). Feedback Consistency Effects in Visual and Auditory Word Recognition: Where Do We Stand After More Than a Decade? *Journal of Experimental Psychology: Learning Memory and Cognition*, *34*(3), 643–661. https://doi.org/10.1037/0278-7393.34.3.643

Chapter 3 – Time course and brain correlates of the Orthographic Consistency Effect

In this chapter, I report the preliminary results of an MEG study. I posed the following research questions:

RQ1 – What is the time course of the OCE? Does it occur before lexical access or after the time point in which the word is accessed in the mental lexicon?

RQ2 – Which brain networks underlie the OCE? The two accounts that I took into consideration are:

- The *restructuring* account, which posits that phonological representations get 'contaminated' with orthographic information during reading acquisition and orthographic information is processed in the left anterior perisylvian regions (e.g., inferior temporal gyrus—IFG; superior temporal gyrus—STG; supramarginal gyrus—SMG) during auditory language processing.
- The *co-activation* account, which posits the involvement of a wider brain network that spans from the left superior temporal gyrus (STG) to the left fusiform gyrus, particularly the visual word form area (VWFA). This implies that there is a bidirectional flow of information between the phonological (STG) and orthographic (VWFA) domains. This means that the phonological representations activated during auditory language processing activate the corresponding orthographic representations that are processed in the VWFA and they, in turn, contribute to auditory language processing.

Thirty Spanish speakers participated in Experiment 3 by completing an auditory LDT with different stimuli than those of Experiment 1 but with the same manipulation and a passive listening task, a more ecological task, while sitting in an MEG scanner. The spatio-temporal sensor analysis shows an enhanced signal for inconsistent words in the 50-160ms time window involving the fronto-temporal sensors in the LDT. This early effect indicates that the activation of orthographic representations of inconsistent phonemes occurs before lexical access. However, no effect was found for pseudowords which goes against my expectations. Finally, no significant difference between inconsistent and consistent stimuli (for both words and pseudowords) was found in the passive listening task. This might indicate that the OCE occurs only under specific conditions during auditory language perception.

3.1 – Theoretical motivation for Experiment 3

Orthographic effects in auditory language perception have been investigated in numerous studies involving different linguistic tasks: auditory lexical decision tasks (e.g., Ziegler & Ferrand, 1998), semantic categorization tasks (e.g., Pattamadilok et al., 2009a), gender categorization tasks (e.g., Peereman et al., 2009), rime detection tasks (e.g., Ziegler et al., 2004), and metaphonological tasks such as phoneme detection (e.g., Frauenfelder et al., 1990). The behavioral measures that these tasks provide, however, represent the "end product" of the involved cognitive processes. Some researchers also claim that listeners strategically rely on orthographic representations when completing, for example, a phoneme detection task (Cutler et al., 2010) or when dealing with specific lexical characteristics (e.g., low word frequency; Damian & Bowers, 2010). Even though Petrova and colleagues (2011) argue that the OCE affects both high- and low-frequency word recognition and is, therefore, an automatic, pre-lexical effect, it should be noted that they found a stronger OCE during low-frequency word recognition. This suggests that listeners rely on sound-to-grapheme mappings more when they perceive low-frequency words. Nevertheless, electrophysiological research seems to provide more solid evidence for the automatic, pre-lexical account. Some EEG studies show that the OCE is pre-lexical as words with inconsistently-spelled rimes generated a larger negativity in the 350ms time-window than words with consistently-spelled rimes (Perre et al., 2009, 2011; Perre & Ziegler, 2008). Perre and Ziegler (2008) especially claim that the OCE occurs online because they found a time-locked ERP response of the OCE deriving from the manipulation of the onset of the inconsistent chunk within their stimuli (i.e., inconsistency in the first or second syllable). These studies, however, did not consider that the manipulation at the syllable/rime level cannot exclude that the OCE occurred after lexical access. In fact, even though the stimuli were always matched on uniqueness point (i.e., the discrimination point at which an auditory stimulus is unequivocally recognized), the OCE found in the late time window (500-600ms after stimulus onset) by Perre and Ziegler (2008) occurred after the complete lexical retrieval of the word which weakens their pre-lexicality argument.

In Experiment 3, I tackled this methodological issue by using inconsistent stimuli whose inconsistency was on the very first phoneme—before the uniqueness point. Experiment 3 also targeted a transparent language (i.e., Spanish) whereas previous neuroimaging studies investigated primarily French, an opaque language. As I argued in Experiment 1, investigating the OCE in a transparent language helps us understand whether literacy affects auditory language perception in general or only when the language has an opaque writing system. Even though Experiment 1 did not show any significant OCE in Spanish pseudoword recognition, I argued that the OCE I found behaviorally in Spanish is pre-lexical because the effect was detected in a language with very few inconsistencies that are almost always governed by orthographic rules. I believe that an online measurement of the effect with MEG would further sustain my argument because the MEG technique allows for a more fine-grained observation. This technique shows on-line neural activity occurring from the presentation of the stimulus to the final lexical decision. For Experiment 3, I therefore expected that the processing of inconsistent stimuli (i.e., both words and pseudowords) would enhance the MEG signal in an early time window (i.e., around the first 100ms from stimulus onset) in an auditory LDT.

Concerning the strategic resort to orthographic information based on specific experimental task requirements, Pattamadilok and colleagues (2014) showed that listeners are affected by orthographic knowledge during inattentive listening. They employed an oddball paradigm that showed a larger mismatch negativity (MMN) for inconsistent words. These findings support the argument that the OCE is pervasive in auditory language perception. However, that study has also the limitation of investigating only words. I argue that the exclusion of pseudowords makes their final claim weaker because an OCE in auditory pseudoword perception would probably be even a stronger piece of evidence of the pervasiveness of the effect in this domain. Therefore, I tackled this issue by running a passive listening task in which participants were asked to attentively listen to the same stimuli of the auditory LDT. I expected a similar pattern of results as those of the LDT. An OCE in passive listening task would imply that the phonology-orthography interaction is automatic and it occurs even during low-level auditory processing and in a more ecological experimental paradigm.

Another open question concerns the language network dynamics underlying the OCE. More specifically, it is still unclear whether an auditory speech input automatically co-activates its orthographic representations or whether the phonemic representation activated during speech processing are "imbued" with orthographic information. As I discussed in Chapter 1, the "co-activation hypothesis" posits that there should be an involvement of the visual word form area (VWFA; especially the fusiform gyrus), a hub for orthographic processing (Tsapkini & Rapp, 2010), jointly with those areas involved in phonological processing. Alternatively, the "restructuring account" posits that the OCE is reflected by a different enhancement of those brain areas related to phonological processing. Investigating which brain correlates underlie the OCE is also important to understand how literacy impacts auditory language perception because both brain network models can be seen as a consequence of literacy.

Regarding this issue, I will present only preliminary results from an analysis that has a more approximate spatial resolution, namely the spatio-temporal sensor analysis. However, finding the cluster of sensors in which the signal enhancement was significant could give an idea of the brain regions involved, bearing in mind that a significant cluster of sensors does not imply that each sensor detected a significant difference in the signal (Sassenhagen & Draschkow, 2019).

3.2 – Methodology

3.2.1 – Participants

Thirty right-handed Spanish speakers (19 female, M_{age} = 22.99 years, *SD* = 2.11 years) participated in Experiment 3 but in no other experiment presented in this thesis. All participants had university-level education; their ages ranged from 18 to 30 years.

Based on self-report, participants had been exposed to their native language since birth and they all predominantly used their native language in everyday life. No specific learning impairments and no hearing or uncorrected vision problems were reported. The participants were recruited from the participant pool of the BCBL (Spain). Participants received monetary compensation for their participation and all signed consent forms, previously approved by the

BCBL's Ethics Committee and by the Bioethics Commission of the University of Barcelona, before starting the experiment.

3.2.2. – Materials

The stimuli were recorded by the same Spanish male speaker and with the same apparatus described in Experiment 1. The stimulus list consisted of 160 words and 160 pseudowords. Like in the other LDT experiments, 50% of the stimuli were consistent while the other half was inconsistent, meaning that at least the first phoneme could be spelled in multiple ways. All inconsistent phonemes were consonants. Unlike Experiment 1, the Spanish stimuli of Experiment 3 were all disyllabic and the mean word frequency was lower (36.92/million vs. 54.73/million). As demonstrated by Petrova and colleagues (2011), the OCE can be detected in both low- and high-frequency words. If anything, lower-frequency words would yield a stronger OCE. All stimuli were matched on the same variables (see Appendix C) as those considered in Experiment 1 across *Consistency* conditions. The variables were also retrieved from the EsPal database (Duchon et al., 2013b) while the phonological and orthographic neighborhood densities of both words and pseudowords were determined with CLEARPOND (Marian et al., 2012a).

3.2.3 – Procedure and MEG data acquisition

The MEG data were acquired in a magnetically shielded room with a whole-scalp system (Elekta Neuromag, Helsinki, Finland) and the bandpass filter set to 0.03–330 Hz, 1 kHz sampling rate. The participants' head position was continuously monitored through five Head Position Indicator (HPI) coils. Three anatomical fiducials (i.e., nasion and left and right prearicular points) plus around 300 additional points registered over the scalp and nose area were digitalized with a 3D digitizer (Fastrak Polhemus, Colchester, VA, USA).

The whole MEG data acquisition consisted of four main blocks, each corresponding to one of four tasks: an auditory lexical decision task, a visual lexical decision task, a passive listening task, and a passive reading task. The first two blocks were always LDTs while the last two were always passive tasks, the order was counterbalanced within the two sub-blocks as well as the handedness of the responses were counterbalanced across participants. The entire experiment had a duration of around 90 minutes. All auditory stimuli were delivered with a random inter-stimulus interval (ISI; from 1.5 to 2s) via nonmagnetic plastic in-ear headphones. In this chapter, I present the preliminary results of the auditory tasks only.

The auditory LDT was administered with Psychotoolbox 3 running on MATLAB 2014b (The MathWorks Inc., Natick, MA, USA). The participants were instructed to respond as quickly and accurately as possible by pressing the respective button for words and for pseudowords. To get familiar with the task, a trial session with 10 stimuli—five words and five pseudowords—was run and feedback about the correctness of the responses was provided only in this section. During the regular task, a fixation cross appeared on the screen for 500ms and then an auditory stimulus was presented through the headphones: the presentation of the stimuli was randomized for each participant. The participants then had 2000ms to respond and a break every 80 trials to avoid fatigue. In total, the LDT lasted around 20 minutes.

The passive listening task was administered with the same software as the auditory LDT. To avoid any repetition effect, the passive listening task was always administered after a visual task (visual LDT or passive reading task) which had different stimuli. The participants were instructed to listen carefully to the stimuli since their attention was going to be tested. This paradigm was considered more ecological than an LDT, but it also has the drawback of not being able of controlling whether the participants a) sustain their attention throughout the task and b) recognize the stimuli correctly. Concerning sustained attention throughout the task, I tried to tackle the issue by inserting a two-back test quite frequently during the passive listening task. The test consisted of asking the participants to recall whether the second-to-last stimulus they listened to so far was a real word or not in Spanish. The response buttons were then the same as those used during the auditory LDT. Concerning the issue of the accurate recognition of the stimuli (i.e., whether words were recognized as words and pseudowords as pseudowords), I relied on the fact that the participants performed at ceiling in the auditory LDT.

3.2.4 – MEG Data Analysis

The MEG data were pre-processed with MaxFilter 2.2 with which signal-noise separation and bad channel removal were performed. To separate external noise from head-internal signal, temporal extension of the signal space separation (Taulu et al., 2005) was applied. Noisy and flat channels were detected automatically, cross-checked manually and subsequently interpolated using a field interpolation method which resulted in using only good channels for interpolation. Next, the data were low-pass filtered at 40 Hz (finite impulse response filter with the hamming window) and eye-blinks, heartbeats, and other artifacts were removed with the independent component analysis (ICA) implemented in MNE Python (Gramfort et al., 2013). Subsequently data was epoched from 0 to 400ms aligned to the stimulus onset, removing noisy epochs with high sensor amplitudes (cut-off thresholds 4000 fT for magnetometers and 4000 fT/cm for gradiometers) and finally data epochs were baseline corrected using the stimulus onset.

For the main analysis, all participants' data was averaged within *Consistency* condition taking words and pseudowords separately and considering only trials with correct responses. Subtractions of interest were performed across all time-points within the epoch and the difference signals for all participants were analyzed with a one-sample spatio-temporal permutation *t*-tests across all time-points and sensors (Maris & Oostenveld, 2007) to identify significant sensors and time-points. This test was performed on RMS (root mean square) combined gradiometer pairs and magnetometers separately. Thus, the contrasts of interest only included the test of the main effect of *Consistency* (inconsistent-consistent).

3.3 – Results of the auditory LDT

In the following section, only the MEG data are presented because only accuracy was recorded as a behavioral measure. Due to technical issues, RTs were not measured during the task. Overall, the participants were at ceiling in the auditory LDT with a mean accuracy of 93.39%. Due to the ceiling effect, no further analysis was possible with accuracy. It should be noted that the data presented in this chapter are preliminary since I only present and discuss the event-related magnetic fields (ERF) data which primarily address the issue of the timing of the OCE. However, I also make some considerations about the brain network underlying the OCE in speech processing based on the spatio-temporal sensor analysis.

The averaged data of 26 participants (four had to be excluded due to extremely noisy data) shows that there is a different pattern in the signal between inconsistent and consistent words at 150ms and at 200ms, as shown in Figure 6.

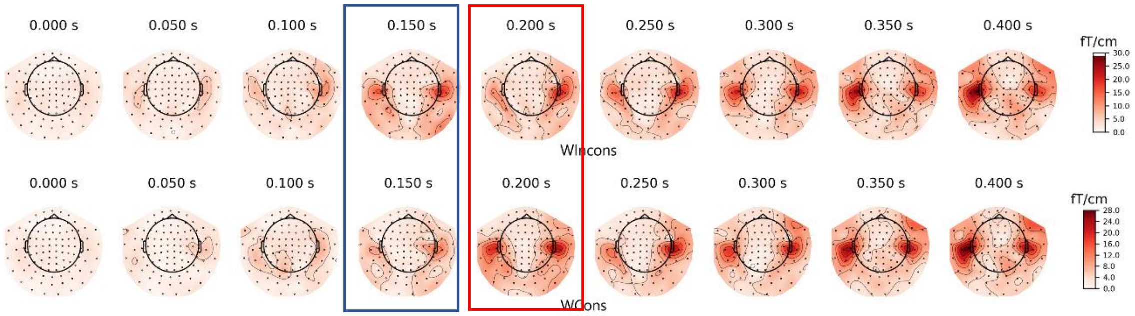


Figure 6. Topography maps based on magnetometers showing the signal enhancement during inconsistent (upper panel) vs. consistent (lower panel) word processing. Darker red coloring means enhanced signal.

The ERF analysis performed by means of the one-sample spatio-temporal *t*-test reveals that the amplitude of the signal of inconsistent words is significantly higher (*p*=0.049) from that of consistent words in the fronto-temporal sensors in the 57-165ms time-window, as shown in Figure 7. This different in the amplitude represents the ERF component for the OCE while the time-window in which it occurs suggests that the orthographic processing runs in parallel with phonological processing

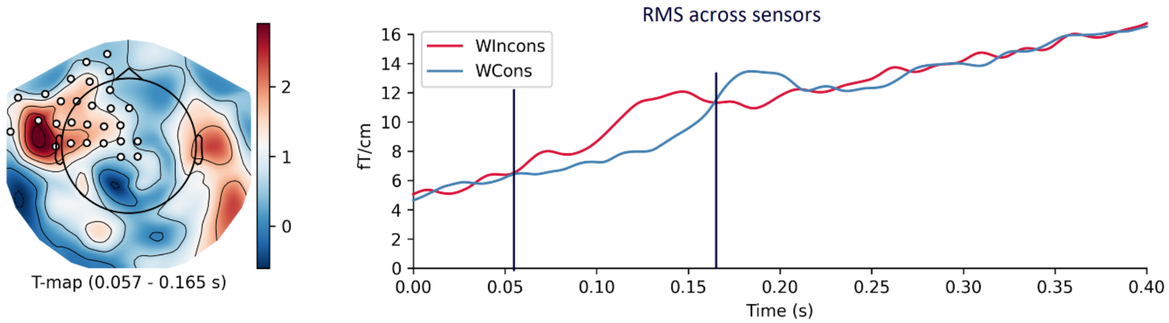


Figure 7. Left—Time-averaged topographical t-value map of the word inconsistent-consistent contrast showing the OCE emerging in the left frontotemporal sensors (marked in white). Right – Root mean square (RMS) signal plotted in the gradiometer which showed the peak of the OCE in the 57 to 165ms time-window.

Conversely, looking at the topographies of the averaged data of inconsistent and consistent pseudowords, it seems that there is no significant difference in the signal between the two conditions, as shown in Figure 8. The observation is statistically confirmed by the one-sample spatio-temporal *t*-test which does not report any significant difference within the *Consistency* condition for pseudowords.

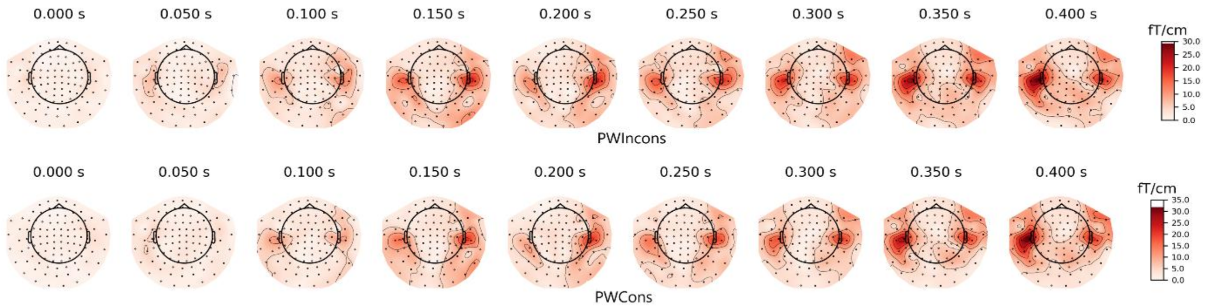


Figure 8. Topography maps based on magnetometers showing the signal enhancement during inconsistent (upper panel) vs. consistent (lower panel) pseudoword processing. Darker red coloring means enhanced signal.

To explain the absence of the OCE in pseudoword processing in the auditory LDT, I further scrutinized the stimuli. Since the auditory stimuli were produced by a human being, I explored whether the duration of the first syllable was significantly different between words and pseudowords. In fact, even though the native Spanish speaker was instructed to produce the pseudowords as word-like as possible, it could be that the speaker hyperarticulated pseudowords. If this was the case, this acoustic cue could be used by the participants of the LDT as a distinctive mark for pseudowords.

The first syllable duration of all stimuli was measured semi-automatically by an external rater who used Praat (Boersma & Weenink, 2022). Overall, the first syllable of words has a mean duration of 278ms (SD = 0.068) whereas the mean duration of the first syllable of pseudowords is 308ms (SD = 0.09). A Welch's *t*-test was run to compare the first syllable durations of the two groups, and it reported a significant difference ($t = 3.38$, $p < 0.001$). This supports the hypothesis of a strategic use of hyperarticulation to discriminate pseudowords from words.

3.4 – Results of the passive listening task

Due to technical problems, the data of only 22 participants could be analyzed. The MEG data analysis was the same as that of the auditory LDT since the same pipeline was employed. Similarly, the same statistical analysis, namely the one-sample spatio-temporal *t*-test, was performed for words and pseudowords separately. In sum, neither words nor pseudowords presented any difference in the *Consistency* condition implying that the OCE occurs under specific conditions during auditory language perception.

The absence of an OCE in word processing could be due to weaker lexical access required by the passive task. It could be, in fact, the lexical threshold is much lower in the passive listening task than in an LDT because there is no stimulus classification during the passive listening task. This might then imply that listeners do not need to rely on orthography when the task demand is low. This argument is similar to what Ventura and colleagues (2004) claim for the absence of the OCE in the shadowing task. However, given that the results of the passive listening task come from only 22 participants, no strong argument can be made about the absence of a significant effect in this more ecological task.

3.5 – General discussion

In Experiment 3, I addressed the still debated issue of the time-course of the OCE and the brain networks underlying it. A further issue that I addressed in this experiment was related to whether the OCE is the result of a strategic use of orthographic information depending on the experimental task at hand or whether it is pervasive in auditory language perception. To this end, Spanish speakers performed an auditory LDT and a passive listening task while their neural activity was recorded in an MEG scanner. Based on previous research (e.g., Perre & Ziegler, 2008), it was expected that the OCE would occur in an early time-window and it would precede lexical access. I expected a similar pattern of results in the passive listening task which would support the argument of the pervasiveness of the OCE in this domain (Pattamadilok et al., 2014). In line with previous research, the auditory LDT showed that inconsistent words generated a larger amplitude of the ERF signal in the early 56-165ms time-window. This means that the OCE precedes lexical access and integration, normally occurring around 400ms after the stimulus onset (Chwilla et al., 1995). Considering that the inconsistent phoneme was always on the first position of the word, the higher amplitude of the signal generated by inconsistent phoneme-to-grapheme mappings in that early time-window provides clear evidence that orthographic processing happens in parallel with phonological processing in auditory word recognition. As the BIA model (Grainger & Ferrand, 1996) posits, an auditory input activates the respective sub-lexical units, such as phonemes, which, at the same time, activate the corresponding orthographic units. Experiment 3 corroborates what was found in Experiment 1 by providing electrophysiological evidence for the pre-lexical nature of the OCE thanks to the online measurements (i.e., ERFs) and the time resolution obtainable with the MEG technique.

The absence of the OCE in auditory pseudoword processing went against my expectations. This pattern (i.e., an OCE for Spanish word but not pseudoword processing) is, however, in line with what was found behaviorally in Experiment 1. On the one hand, the replication of the findings in two different experiments with different materials and participants reassures their solidity. On the other hand, what happens during pseudoword processing and, consequently, the role of orthography in auditory language perception remain unclear. One possibility is that the kind of task, namely the auditory LDT, does not enable orthography to influence pseudoword processing because pseudoword are classified with “no” responses without further analysis in the LDT. It could be that “no” responses in an LDT are given when no threshold for lexical access is reached by pseudowords (Ziegler & Ferrand, 1998) basically washing out any orthographic effect. The “time-out mechanism” (Taft, 2011) is also supported by the longer RTs for pseudowords than for words found in Experiment 1 yet only for Spanish, since an OCE was found in French pseudoword processing. Nevertheless, I further scrutinized the experimental stimuli to understand whether there could be another explanation for this null effect. It appears that Spanish pseudowords were hyperarticulated by the Spanish native speaker who recorded them. Probably, no OCE in auditory pseudoword processing was found because the participants could have strategically used hyperarticulation as a cue to identify pseudowords and this could have cancelled out the effect at study.

The null results of the passive listening task also went against my hypotheses, although limited conclusions can be drawn from null effects. One should consider that the sample size was smaller than that of the auditory LDT which showed a significant OCE in auditory word perception, yet quite close to the significance threshold (*p*=.049). Therefore, I cannot exclude that the results of the passive listening task could be related to lower statistical power. To this end, it is my intention to further analyze the MEG dataset that I have: I will attempt to recover some excluded participants with more advanced data preprocessing. In fact, I would like to reiterate that these results are preliminary.

Concerning the brain network that underlies the phonology-orthography interplay in speech perception, the present analysis allows to cautiously argue that phonemic representations are 'contaminated' with orthographic information (see e.g., Montant et al., 2011). This argument is based on the finding that the higher amplitude generated by inconsistent words was detected in a cluster of fronto-temporal sensors which can be roughly matched with the perisylvian regions typically attributed to phonological processing. From a more theoretical perspective, Experiment 3 provides clear evidence that alphabetic literacy acquisition not only establishes a relationship between sounds and letters, but it also reshapes already existing phonemic representations.

To conclude, Experiment 3 shows that the OCE influences word recognition of Spanish speakers at the pre-lexical level. This effect, however, seems to emerge when the word needs to be fully accessed, implying that the orthographic information becomes relevant only when the cognitive cost required by a linguistic task is higher. One could argue that the phonemic representations activated during the LDT and those activated during the passive listening task are not the same. There are indeed in some accounts in the literature that posit that there are multiple phonological representations: a type is more speech-based (i.e., acoustic) while the other type is more abstract (see e.g, Friedrich, 1990). It could be that the LDT activates more abstract phonological representations, which also contain orthographic information. Conversely, a passive listening task, which requires more low-level auditory processing, activates more speech-based representations that do not contain orthographic representations. Concerning the brain network that underlies the phonology-orthography interface during speech perception, the spatio-temporal sensor analysis suggests that brain regions typically attributed to phonological processing (e.g., IFG) are also involved in some sort of orthographic processing during speech processing. However, source-reconstruction and connectivity analyses which combine both MEG data and T1 MRI scans will provide more detailed, fine-grained insights regarding this issue.

References

Boersma, P., & Weenink, D. (2022). *Praat: Doing phonetics by computer* (6.2.12) [Computer software]. http://www.praat.org/

Chwilla, D. J., Brown, C. M., & Hagoort, P. (1995). The N400 as a function of the level of processing. *Psychophysiology*, *32*(3), 274–285. https://doi.org/10.1111/j.1469-8986.1995.tb02956.x

Cutler, A., Treiman, R., & van Ooijen, B. (2010). Strategic Deployment of Orthographic Knowledge in Phoneme Detection. *Language and Speech*, *53*(3), 307–320. https://doi.org/10.1177/0023830910371445

Damian, M. F., & Bowers, J. S. (2010). Orthographic effects in rhyme monitoring tasks: Are they automatic? *European Journal of Cognitive Psychology*, *22*(1), 106–116. https://doi.org/10.1080/09541440902734263

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

Frauenfelder, U. H., Segui, J., & Dijkstra, T. (1990). Lexical Effects in Phonemic Processing: Facilitatory or Inhibitory? *Journal of Experimental Psychology: Human Perception and Performance*, *16*(1), 77–91. https://doi.org/10.1037/0096-1523.16.1.77

Friedrich, F. J. (1990). Multiple phonological representations and verbal short-term memory. In G. Vallar & T. Shallice (Eds.), *Neuropsychological Impairments of Short-Term Memory* (pp. 74–93). Cambridge University Press. https://doi.org/10.1017/CBO9780511665547.005

Grainger, J., & Ferrand, L. (1996). Masked orthographic and phonological priming in visual word recognition and naming: Cross-task comparisons. *Journal of Memory and Language*, *35*(5), 623–647. https://doi.org/10.1006/jmla.1996.0033

Gramfort, A., Luessi, M., Larson, E., Engemann, D., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L., & Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience*, *7*, 1–13. https://doi.org/10.3389/fnins.2013.00267

Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). CLEARPOND: Cross-Linguistic Easy-Access Resource for Phonological and Orthographic Neighborhood Densities. *PLoS ONE*, *7*(8), e43230. https://doi.org/10.1371/journal.pone.0043230

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Montant, M., Schön, D., Anton, J. L., & Ziegler, J. C. (2011). Orthographic contamination of Broca's area. *Frontiers in Psychology*, *2*(DEC), 1–10. https://doi.org/10.3389/fpsyg.2011.00378

Pattamadilok, C., Morais, J., Colin, C., & Kolinsky, R. (2014). Unattentive speech processing is influenced by orthographic knowledge: Evidence from mismatch negativity. *Brain and Language*, *137*, 103–111. https://doi.org/10.1016/j.bandl.2014.08.005

Pattamadilok, C., Perre, L., Dufau, S., & Ziegler, J. C. (2009). On-line Orthographic Influences on Spoken Language in a Semantic Task. *Journal of Cognitive Neuroscience*, *21*(1), 169–179. https://doi.org/10.1162/jocn.2009.21014

Peereman, R., Dufour, S., & Burt, J. S. (2009). Orthographic influences in spoken word recognition: The consistency effect in semantic and gender categorization tasks. *Psychonomic Bulletin & Review*, *16*(2), 363–368. https://doi.org/10.3758/PBR.16.2.363

Perre, L., Bertrand, D., & Ziegler, J. C. (2011). Literacy affects spoken language in a non-linguistic task: An ERP study. *Frontiers in Psychology*, *2*(OCT), 1–8. https://doi.org/10.3389/fpsyg.2011.00274

Perre, L., Pattamadilok, C., Montant, M., & Ziegler, J. C. (2009). Orthographic effects in spoken language: On-line activation or phonological restructuring? *Brain Research*, *1275*, 73–80. https://doi.org/10.1016/j.brainres.2009.04.018

Perre, L., & Ziegler, J. C. (2008). On-line activation of orthography in spoken word recognition. *Brain Research*. https://doi.org/10.1016/j.brainres.2007.10.084

Petrova, A., Gaskell, M. G., & Ferrand, L. (2011). Orthographic consistency and word-frequency effects in auditory word recognition: New evidence from lexical decision and rime detection. *Frontiers in Psychology*, *2*(OCT), 1–11. https://doi.org/10.3389/fpsyg.2011.00263

Sassenhagen, J., & Draschkow, D. (2019). Cluster-based permutation tests of MEG/EEG data do not establish significance of effect latency or location. *Psychophysiology*, *56*(6), e13335. https://doi.org/10.1111/psyp.13335

Taft, M. (2011). Orthographic influences when processing spoken pseudowords: Theoretical implications. *Frontiers in Psychology*, *2*(JUN), 1–7. https://doi.org/10.3389/fpsyg.2011.00140

Taulu, S., Simola, J., & Kajola, M. (2005). Applications of the signal space separation method. *IEEE Transactions on Signal Processing*, *53*(9), 3359–3372. https://doi.org/10.1109/TSP.2005.853302

The MathWorks Inc. (n.d.). *MATLAB* (8.4.0 150421) [Computer software]. The MathWorks Inc.

Tsapkini, K., & Rapp, B. (2010). The orthography-specific functions of the left fusiform gyrus: Evidence of modality and category specificity. *Cortex*, *46*(2), 185–205. https://doi.org/10.1016/j.cortex.2009.02.025

Ventura, P., Morais, J., Pattamadilok, C., & Kolinsky, R. (2004). The locus of the orthographic consistency effect in auditory word recognition. *Language and Cognitive Processes*, *19*(1), 57–95. https://doi.org/10.1080/01690960344000134

Ziegler, J. C., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin and Review*, *5*(4), 683–689. https://doi.org/10.3758/BF03208845

Ziegler, J. C., Ferrand, L., & Montant, M. (2004). Visual phonology: The effects of orthographic consistency on different auditory word recognition tasks. *Memory and Cognition*, *32*(5), 732–741. https://doi.org/10.3758/BF03195863

Chapter 4 – The influence of literacy on speech production

> In the present chapter, I present two experiments that explored whether and to what extent literacy influences language production. The Orthographic Consistency Effect (OCE) in language production has not been much investigated, especially cross-linguistically, and the results reported so far are contradictory since some studies found it and some others did not. Experiments 4 (adults) and 5 (children) consisted of a reading aloud task (RAT) and a picture naming task (PNT) in which the same participants as in Chapter 2 took part. The manipulation of the stimuli was always at the phonemic level. In the RAT, both words and pseudowords could be either consistent (i.e., with phonemes with only one corresponding grapheme) or inconsistent (i.e., with phonemes with multiple corresponding graphemes). In the PNT, the pictures' referents were either orthographically consistent or inconsistent. The expectation was to find an OCE especially in pseudoword production. Overall, the result showed that Spanish-speaking seven-year-olds were faster at producing consistent items in the RAT but not in the PNT. However, Spanish-speaking adults did not appear to be affected by orthographic (in)consistency in speech production.

4.1 – Theoretical motivations for Experiment 4

As already described in Chapter 1, there are only a few studies investigating the OCE in language production. Damian and Bowers (2003) employed the form-preparation paradigm (A. S. Meyer, 1990) to test the influence of phoneme-to-grapheme mappings on language production in English. The paradigm requires participants to produce the last word of a previously learned word pair. If the target word started with the same phoneme as the other word in the pair, but the phoneme had different spellings (e.g., 'c̲offee' and '*k̲ennel*'), participants needed more time to respond. The results were interpreted in the light of interaction models in reading aloud like the dual-route cascaded model (DRC; Coltheart et al., 2001b; Coltheart & Rastle, 1994): a series of orthographic representations are activated from reading and they consequently activate the corresponding phonological representations that will eventually be translated into articulatory gestures. In turn, the activated phonological representations activate all the relative orthographic representations. If the phoneme is consistent, then it would *feed back* to the already activated correspondent grapheme. If the phoneme is inconsistent, it would then co-activate other graphemes in addition to the already activated one. Damian and Bowers cautiously conclude that consistent phoneme-to-grapheme mappings, as a by-product of literacy, facilitate language production. In other terms, they found an OCE in language production as English inconsistent words co-activated all the possible orthographic representations of an inconsistent phoneme which slowed down word production.

However, Damian and Bowers' (year) findings were not consistently replicated in studies investigating the OCE in language production. In a study on Dutch, Roelofs (2006) did not find any difference in production latencies between consistent and inconsistent words even though he employed the same paradigm as Damian and Bowers (2003) which did not involve any orthography. Roelofs suggested that the discrepancy with Damian and Bowers depended on the target languages at study: Roelofs investigated Dutch, a transparent language, while Damian and Bowers investigated English, a highly opaque language. He concluded that there might be a different degree of interaction between orthography and phonology in language production which depends on the opacity of the writing system. In his study, however, Roelofs also found an OCE in word reading and he concluded that orthography plays a role in language production only when there are specific task demands, such as reading aloud a written word. In the reading aloud paradigm, orthography is relevant and therefore the OCE could be detected. Conversely, orthography does not appear to affect speech production latencies in tasks such as picture naming where no orthographic forms are presented to the participants.

The first aim of Experiment 4 is to systematically compare the role of phoneme-to-grapheme mappings in language production in French, as a representative of opaque languages, and Spanish, as a representative of transparent languages, in the reading aloud paradigm. The fact that orthography is present in the experimental paradigm can be seen as an advantage in the investigation of OCE in language production. Roelofs' argument on the different impacts of phoneme-to-grapheme mappings on language production assumes that the opacity of the language in question plays a role. This was based on a comparison between English (i.e., the study of Damian and Bowers, 2003) and Dutch (i.e., his study). On the one hand, English is a language with many inconsistencies in both phoneme-to-grapheme and grapheme-to-phoneme mappings whereas Dutch is more consistent in grapheme-to-phoneme relation than in phoneme-to-grapheme mappings (Borgwaldt et al., 2010). On the other hand, French is quite inconsistent in phoneme-to-grapheme mappings but rather consistent in grapheme-to-phoneme mappings (Peereman & Content, 1999). Similarly, Spanish has more inconsistencies in phoneme-to-grapheme mappings than in grapheme-to-phoneme mappings which made the French-Spanish comparison more appropriate to investigate the OCE in language production since the effect is based on sound-to-spelling mappings. Moreover, both French and Spanish are languages with simple syllabic structures (Seymour, Aro, & Erskine, 2003). According to their cross-linguistic study, Seymour and colleagues (2003) found that languages with complex syllabic structures (e.g., English, Danish, Dutch) require different strategies in decoding, especially during reading acquisition, resulting in a significant disadvantage in reading simple pseudowords in comparison to speakers of languages with simple syllabic structures.

Consequently, I argue that the French-Spanish comparison was better suited to investigate the OCE in language production since there would not be any interference from the syllabic structures involved. It should be noted, however, that the disadvantage found by Seymour and colleagues was only limited to pseudoword and not to word reading. Therefore, I expected a magnified OCE in pseudoword reading compared to word reading, resulting in longer latencies for inconsistent than consistent pseudowords. Concerning the cross-linguistic differences, I expected that Spanish could also show an OCE in word reading because Spanish readers could rely more on the sub-lexical route during word reading than French readers. This expectation is based on previous literature (e.g., Ziegler & Goswami, 2005) but also on the results of Experiment 1 where Spanish listeners showed a clear OCE in auditory language processing.

The second aim of Experiment 4 is to investigate the OCE in a language production paradigm in which there is no orthography involved like picture naming. This paradigm has the advantage of initiating the production process from semantics and then progressively involving phonological and orthographic representations (see Levelt et al., 1991 for an overview). Moreover, this paradigm has not been often employed in previous research and the results are not consistent. Both Roelofs (2006) and Alario and colleagues (2007) did not find an OCE in language production with an overt picture naming paradigm in Dutch and French respectively. Yet, Qu and Damian (2019) found strong evidence for orthographic influence in language production in Chinese using a picture naming paradigm. Similarly, Rastle and colleagues (2011) found an OCE in language production in a picture naming task in which the participants had to name pictures depicting novel words in English. The novelty of the second part of Experiment 4 was to systematically compare an opaque language like French and a transparent language like Spanish in a picture naming paradigm. The cross-linguistic approach was important to explore possible differences deriving from the opacity of the language as was argued by Roelofs (2006) in language production experiments employing other paradigms. In this experiment, the two following hypotheses are thus at test:

a) The OCE appears also in picture naming indicating that inconsistencies in phoneme-to-grapheme mappings affect language production not only when orthography is relevant (i.e., present) in the task. Cross-linguistically, Spanish should show a magnified OCE in comparison to French for the same reasons mentioned in the RAT.
b) No OCE is present in PNT because orthography is not relevant for the task (Roelofs, 2006). Alternatively, another argument is that orthography might play a role not in language production per se but rather in memorization processes in word-association procedures (Alario et al., 2007). In this case, no cross-linguistic difference is expected because this output would imply that the kind of task simply does not involve orthographic representations linked to the phonemic representations.

4.2. – Methodology

Experiment 4 consisted of two tasks: a reading aloud task (RAT) and an overt picture naming task (PNT). During the RAT, participants saw a visual stimulus, either a word or a pseudoword, and they had to read it out loud as quickly and accurately as possible. During the PNT, participants saw pictures depicting real concepts and they had to name them as quickly and accurately as possible.

4.2.1 – Participants

The same 30 L1-Spanish and 30 L1-French of Experiment 1 participated in Experiment 4.

4.2.2 – Materials

The stimuli list of the RAT consisted of 40 words and 40 pseudowords for each language. Like in previous experiments, pseudowords were created by modifying the first phoneme of each syllable, always respecting the languages' phonotactic rules. The type of manipulation was the same as that of previous experiments: 50% of the stimuli were consistent and the remaining 50% were inconsistent at least on the first phoneme, which was always a consonant. The stimuli were matched on a series of confounding variables within languages and across conditions. Across languages, the stimuli were matched on variables such as word frequency (Zipf's), number of letters, orthographic neighbors (see Appendix D).

Concerning the stimuli list of the PNT, 60 pictures for each language were selected from the MultiPic database (Duñabeitia et al., 2018) when available (115). Five pictures were retrieved from royalty-free online sources (2 Spanish and 3 French). The MultiPic database offers stimuli in several languages, including French and Spanish, which are normed on visual complexity and naming labels. The pictures consisted of colored drawings representing real referents. For each language, 50% of the words represented by the pictures were consistent and the remaining 50% were inconsistent. The manipulation was again on at least the first phoneme of the inconsistent words. The stimuli matching was also analogous to that of the other stimuli lists employed in the other experiments (see Appendix F and G).

It should be noted that the words of the RAT and the PNT were different.

4.2.3 – Apparatus and procedure

The apparatus and the setting of Experiment 4 were the same as the one described in Experiment 1 since the participants were tested in a single experimental session. The headphones had an external microphone for voice recording.

In the RAT, each trial started with a fixation cross which appeared in the center of the screen for 500ms. Afterwards, the visual stimulus was presented in the center of the screen for 3000ms, and the participants had to read it out loud. The stimuli were presented in a randomized order for each participant. Before the next trial started, there was an inter-stimulus interval (ISI) of 500ms. The background of the experiment was white, and the visual stimuli were presented in black Droid Sans Mono font (font-size 32px). The voice recording started as soon as the stimulus was displayed on the screen. The task lasted around 10 minutes.

In the PNT, the participants saw first all the pictures during a familiarization session in which the experimenter asked the participants to name them one after the other and to rate them on their visual complexity. If a picture was named with a word that was not intended for the experimental purpose (e.g., if a picture of a dog was named 'animal'), the experimenter would correct the participant and ask the participant to use the correct word during the actual task. The visual complexity rating was administered to get the rating values also for the five pictures that were not taken from the MultiPic database. The rating was given based on a 5-point Likert scale in which 1 referred to 'very simple' and 5 to 'very complex'. The phrasing of instructions for the rating task was the same as the one used by Duñabeitia and colleagues (2018) for the MultiPic project. In the PNT trials started with a fixation cross was displayed in the center of the screen with white background for 500ms and then a picture (300x300px) was displayed in the center of the screen for 3000ms. The ISI lasted 500ms. The voice recording started when the stimulus appeared on the screen. The presentation of the stimuli was randomized across participants and no feedback was provided during the task. The task lasted around 20 minutes.

4.3 – Results

The measured variable of both RAT and PNT was speech onset-time (SOT). SOT represents the time delay between the stimulus onset and the acoustic onset of the response articulation. Longer SOTs would be interpreted as more effortful speech programming. Two raters measured SOT in Praat software (Boersma & Weenink, 2022) and each rater checked the measurements of the other. The utterances were coded as correct only if they were produced without any hesitation, discernibly, and accurately.

4.3.1 – Results of the reading-aloud task (adults)

Due to a technical problem, the data of a Spanish participant was not recorded during the task. Overall, both French and Spanish participants performed at ceiling (French: 96.88%; Spanish: 96.42%). The incorrect responses represent 3.35% of the entire dataset and they were excluded from the analysis. Only SOTs above and below 3SD from the mean were removed on a by-participant basis (1.89% of correct responses) since there were not any extremely long (>2000ms) or extremely short responses (>150ms).

Overall, both French and Spanish present on average longer SOTs for inconsistent pseudowords than for consistent pseudowords. As for words, there was not such numerical trend. Table 8 summarizes the descriptive statistics of the RAT.

Table 8. Mean SOT in ms of the RAT for both languages by lexicality and consistency (SD in ms).

	French		Spanish	
	Word	Pseudoword	Word	Pseudoword
Consistent	576 (105)	659 (170)	574 (113)	655 (153)
Inconsistent	575 (112)	690 (193)	570 (106)	687 (168)
Δ	-1	31	-4	32

Note: Delta expresses the difference in milliseconds between inconsistent and consistent items. SD in parentheses.

Like in Experiment 1, the main analysis was run in RStudio (version 1.3.1073; RSTudio Team, 2020) using the *lme4* package (Bates et al., 2015). SOT data were analyzed using a generalized linear mixed-effects model (GLMM) with the *glmer* function of *lme4*. The GLMM was run based on the assumption of a Gamma distribution of the data with an identity link (Lo & Andrews, 2015b). The predicted variable was *SOT* (expressed in ms) and the predictors were *Consistency* (consistent= -1, inconsistent=1), *Lexicality* (word=-1, pseudoword=1), and *Language* (French=-1, Spanish=1) which were contrast coded. The three predictors were linked with a three-way interaction term including lower-level interactions. Random intercepts for *Participant* and *Stimulus* with by-*Participant* random slopes for *Lexicality* and *Consistency* were also included. The optimizer BOBYQA was applied to the model to solve convergence problems (Powell, 2009).

The model detected a significant effect for *Lexicality* while *Consistency* was slightly above the significance threshold as shown in Table 9.

Table 9. Main GLMM output from the SOT analysis.

Fixed effects		Estimate	SE	*t*	*Pr (>|t|)*
Intercept		667.04	5.83	114.39	<.001
Consistency		9.10	5.37	1.70	*.0898*
Lexicality		65.67	5.68	11.57	<.001
Language		-3.44	6.97	-0.49	.622
Consistency*Lexicality		6.47	5.29	1.22	.221
Consistency*Language		-1.81	5.30	-0.34	.733
Lexicality*Language		-2.49	6.64	-0.38	.708
Consistency*Lexicality*Language		-1.22	5.21	-0.24	.814
Random effects	*Group*	*Variance*	*SD*	*Correlation*	
Item	Intercept	1.457e+03	38.17		
Participant	Intercept	1.308e+03	36.17		
	Lexicality	3.410e+02	18.47	0.47	
	Consistency	5.738e+01	7.57	0.22	0.16

Note: SE= standard error; Pr (>|t|) p-value calculated from the t-value with lmerTest package (Kuznetsova et al., 2017)*; SD = standard deviation.*

Given that no interaction was found from this statistical analysis, it was not warranted to further investigate the numerical trend for pseudoword production found in both French and Spanish pseudowords.

4.3.2 – Results of the picture naming task (adults)

Both French and Spanish participants performed the task at ceiling with very few errors: both languages had an error rate of 1.39% each.

Only SOTs above and below 3SD from the mean were removed on a by-participant basis (1.01% of correct responses) since there were not any extremely long (>2000ms) or extremely short responses (>150ms).

From a descriptive point of view, there is no apparent trend indicating that inconsistent words were produced more slowly than consistent words in either language as Table 10 shows.

Table 10. Mean SOT in ms of the PNT for both languages by consistency (SD in ms).

	French	Spanish
Consistent	967 (36)	964 (34)
Inconsistent	966 (39)	961 (34)
Δ	-1	-3

Note: Delta expresses the difference in milliseconds between inconsistent and consistent items. SD in parentheses

The statistical analysis was based on the same conditions as those of the RAT with the only exception that the GLMM did not include *Lexicality* as a predictor because the PNT did not contain any pseudowords. As a reminder, the predicted variable was *SOT* (expressed in ms) and the predictors were *Consistency* (consistent= -1, inconsistent=1) and *Language* (French=-1, Spanish=1) which were contrast coded. The two predictors were linked with a simple interaction term. Random intercepts for *Participant* and *Stimulus* with a by-*Participant* random slope for *Consistency* were also included.

Overall, the model did not show any significant effects or interactions (see Appendix X for model output).

4.3.2.1 – Further exploration of the PNT data (adults)

There are important variables that can affect naming speed and naming accuracy such as imageability (i.e., how easily a word evokes the mental image of the referent), familiarity (here intended as subjective word frequency), concept concreteness (i.e., how much the referent of a word can be perceived with the senses) and age of acquisition (AoA; i.e., self-reported estimation of the age at which a word was acquired; see Perret & Bonin, 2019 for an overview on the effect of these variables).

A posteriori, subjective ratings on imageability, familiarity, concreteness, and age of acquisition were collected in an online survey in which 30 L1-French and 30 L1-Spanish speakers participated. These participants did not participate in any other study presented in this thesis. The subjective ratings were collected for a total of 400 words per language also including the words of the stimuli list of the PNT. Imageability, familiarity, and concreteness ratings were given on a 7-point Likert scale whereas AoA ratings were given by indicating age intervals from 0 to 7 years of age.

Overall, the means of all variables of the wordlist used in the PNT were matched across conditions and languages. I decided, however, to further explore the PNT data to see whether these variables are a predictor interacting with *Consistency*. For this purpose, I ran four GLMMs to test each variable separately. I then compared the four models with the *compare_performance* function of the *performance* R package (Lüdecke et al., 2021) to check which variable would better explain the data variance and possibly co-vary in the *Consistency* condition. The analysis showed that the model that better performed in terms of convergence and explained data variability was the one that included *Imageability* as a co-variate. The model shows no significant effects or interactions.

4.4 – Summary of Experiment 4

Experiment 4 showed no evidence for the OCE in language production regardless of the opacity of the writing system and regardless of the relevance of orthography for the task at hand. In fact, in contrast to Roelofs (2006), there was only a numerical tendency for both French and Spanish pseudowords in the reading aloud task indicating that inconsistent pseudowords are more effortful to read aloud than consistent pseudowords. This tendency might underlie decoding strategies in pseudoword reading which, however, are not as strongly impacted by orthographic inconsistencies as in other language processing domains (e.g., auditory language processing or phonological awareness). The discrepancy between these findings and those of Roelofs (2006) may be reconducted to the syllabic structures of the languages in the study. Dutch, compared to French and Spanish, has a more complex syllabic structure. As pointed out by Seymour and colleagues (2003), syllabic complexity determines decoding strategies to the extent that speakers of languages with higher syllabic complexity produce inconsistent pseudowords more effortfully than speakers of languages with simple syllabic structures. It should be noted, however, that Roelofs (2006) found an OCE only in word reading aloud since there were no pseudowords in his experiment. More research is needed to understand the role of syllabic complexity in the OCE in language production.

Both the RAT and the PNT suggested that phoneme-to-grapheme mappings are not involved in word production. This finding is in line with what Alario and colleagues (2007) found in French. As Alario and colleagues argue, the OCE in Damian and Bowers (2003) could be related to the strategic use of orthographic information in the word memorization procedure that the experimental paradigm (i.e., form-preparation paradigm) required. Alternatively, a possible explanation is that no OCE was found in word production because words in both RAT and PNT were high-frequency and, therefore, processed mainly through the lexical route (Barca et al., 2002).

To conclude, Experiment 4 seems to indicate that phoneme-to-grapheme mappings are not automatically involved during language production but rather employed strategically depending on the task at hand. The reason behind this could be that adult speakers have accumulated such language experience that they can read and speak in a quick and 'optimized' way. This argument can be further supported by exploring speakers with smaller linguistic experience-children- perform the same tasks in Experiment 4. My attempt in Experiment 5 is, thus, to address this issue.

4.5 – Theoretical motivation for Experiment 5

As previously mentioned, learning to read in an alphabetic language implies establishing connections between already existing phonological representations with new, less stable, orthographic representations. In Chapter 2, Experiment 2, Spanish seven-year-olds showed an OCE in language perception since they were more accurate at recognizing consistent than inconsistent words and pseudowords. My main argument was that children rely on phonological decoding during language perception which results in the co-activation of orthographic representations of the respective phonemes.

The pervasiveness of the OCE in children's language perception raises therefore the question of whether it is also present in children's language production. The first aim of Experiment 5 was to answer this question. From Experiment 4, I concluded that phoneme-to-grapheme mappings might come at play in adults' language production only in specific contexts (i.e., pseudoword reading), as Alario and colleagues (2007) also argued, even though this is based only on a numerical trend. However, I did not expect the same pattern in children because children do not yet have stable phoneme-to-grapheme and grapheme-to-phoneme mappings. This implies that children are probably more prone to employ phonological decoding rather than other processing strategies. Moreover, children, especially those with a transparent L1, tend to rely more on the sub-lexical route when they learn to read (Ziegler & Goswami, 2005), and they should be consequently more sensitive to the OCE at the phonemic level. Taking all this into account, I expected that children would show an OCE in the RAT for both words and pseudowords. From a cross-linguistic perspective, I expected a smaller language group difference because both French and Spanish children could resort to the sub-lexical route to a more similar extent than their adult counterparts. Yet, as in Experiment 2, this cross-linguistic comparison could not be made due to the COVID-19 pandemic which prevented me from collecting data in French schools due to the enforced health measures.

Concerning the PNT, the two same hypotheses of Experiment 4 were tested. Yet, I expected that children would show an OCE in this task as well because the lemma retrieval would require a stronger orthographic involvement than in adults. This hypothesis is based on the findings of previous studies with dyslexic children who performed more poorly than their control counterparts in PNT, implying an orthographic involvement in the task (Goswami et al., 1999; Swan & Goswami, 1997).

4.6 – Methodology

Like Experiment 4, Experiment 5 also consisted of a RAT and a PNT.

4.6.1 – Participants

The same participants of Experiment 2, that is 45 L1-Spanish seven-year-olds, participated in Experiment 5.

4.6.2 – Materials

Concerning the RAT, the Spanish stimuli list used in Experiment 4 was shortened to 20 words and 20 pseudowords to avoid any fatigue effect in the children. 50% of the stimuli were consistent and 50% were inconsistent. As for the PNT, the same Spanish stimuli list was employed (i.e., 60 pictures with 50% of the linguistic referent consistent and 50% inconsistent).

4.6.3 – Apparatus and procedure

The apparatus and the experimental setting of Experiment 5 were the same as the one described in Experiment 2.

In the RAT, the experimenter explained the task to the child and invited them to read as quickly and accurately as possible what was going to appear at the center of the screen. Like the RAT with adults, a trial consisted of a fixation cross at the center of the screen which would remain for 500ms, then the presentation of the stimulus which, in this case, would remain on the screen until the experimenter pressed a key to continue to next trial. The technical characteristics of the visual stimulus (font, font size, background, etc.) were the same as those of the RAT in Experiment 4. No feedback was provided to the children during the task.

Like Experiment 4, the PNT of Experiment 5 was preceded by a familiarization session in which children named all the pictures once and if the labels they used were not congruent with the experimental word, the experimenter would instruct the children to name that picture accordingly during the main task. The PNT procedure was the same as in the RAT. The experimenter managed the progression of the task so that children could take breaks whenever needed.

4.7. – Results

The measured variables of the RAT were SOT and accuracy whereas only SOT was taken into account for the PNT. The SOT measurement procedure was the same as in Experiment 4.

4.7.1 – Results of the reading-aloud task (children)

Overall, the L1-Spanish seven-year-olds were quite accurate in the RAT as their accuracy rate was 90.5%. From a descriptive point of view, the participants were more accurate at producing consistent stimuli than inconsistent stimuli, as Figure 9 shows.

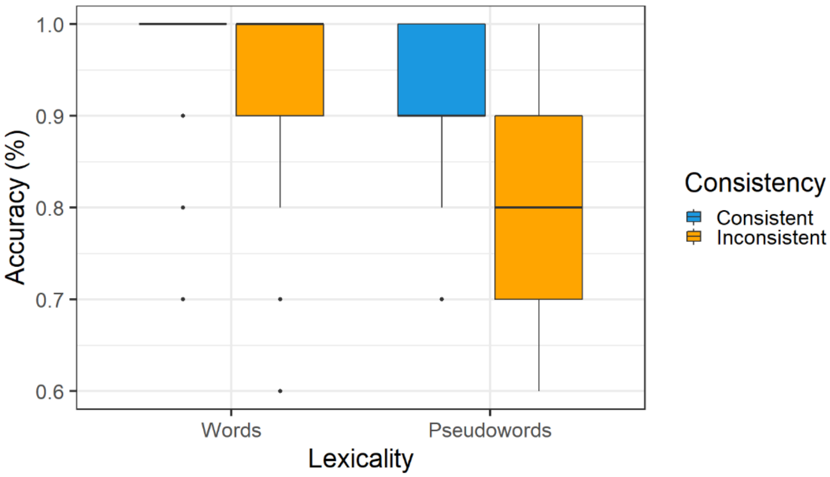


Figure 9. Distribution of accuracy (grouped by participants) across conditions.

The statistical analysis for accuracy was the same as the one employed in Experiment 2 in the LDT: a logistic mixed-effects model with *Accuracy* (correct production = 1, incorrect production = 0) as the dependent variable with fixed effects for *Consistency* (consistent = -1, inconsistent = 1) and *Lexicality* (word = -1, pseudoword = 1) with a simple interaction term. The model also included random intercepts for *Subjects* and Items, as well as by-subject random slopes for *Consistency* and *Lexicality*. Also, in this case, the optimizer BOBYQA was applied to the model to solve convergence problems.

The logistic model reports a main effect for *Consistency* and *Lexicality* but no interaction between the two predictors. Table 11 summarizes the model output.

Table 11. Main logistic mixed-effects model output from the accuracy analysis.

Fixed effects		β	SE	*z*	*Pr (>|z|)*
Intercept		2.77	0.21	13.20	<.001
Consistency		-0.41	0.17	-2.36	.0185
Lexicality		-0.70	0.19	-3.70	.0002
Consistency*Lexicality		0.06	0.17	0.35	.0730
Random effects	*Group*	*Variance*	*SD*	*Correlation*	
Item	Intercept	0.74	0.86		
Participant	Intercept	0.55	0.74		
	Consistency	0.01	0.09	-0.37	
	Lexicality	0.20	0.45	-0.67	-0.45

Note: SE= standard error; Pr (>|z|) p-value calculated from the z-value; SD = standard deviation.

The SOT data trimming was operated according to the same criteria of Experiment 4: removal of excessively long (> 3000ms) and short trials (< 150ms), which represent 7.74% of the correct responses. On a by-participant basis, trials above and below 3SD of the mean RT were also excluded from the analysis (5.1% of the correct responses). Overall, inconsistent stimuli had longer SOTs than consistent stimuli, as summarized in Table 12.

Table 12. Mean SOT of the RAT by lexicality and consistency (SOT in milliseconds).

	Word	Pseudoword
Consistent	1148 (493)	1362 (497)
Inconsistent	1204 (508)	1472 (528)
Δ	56	110

Note: Delta expresses the difference in milliseconds between inconsistent and consistent items. SD in parentheses.

The statistical analysis for the SOT data involved the same GLMM model used in Experiment 4 which was run based on the assumption of a Gamma distribution of the data with an identity link. The predicted variable was *SOT* (expressed in ms) and the predictors were *Consistency* (consistent = -1, inconsistent = 1) and *Lexicality* (word = -1, pseudoword = 1) which were contrast coded. The two predictors were linked with a simple interaction term. Random intercepts for *Participant* and *Stimulus* with by-*Participant* random slopes for *Lexicality* and *Consistency* were also included. Also, in this case, the optimizer BOBYQA was applied to the model to solve convergence problems.

The GLMM reported two strong main effects for *Consistency* and *Lexicality* and no interaction between them. The model output is summarized in Table 13.

Table 13. Main GLMM output from the RT analysis.

Fixed effects	β	SE	*t*	*Pr (>|t|)*
Intercept	1477.65	12.97	113.89	<.001
Consistency	54.28	11.20	4.85	.<.001
Lexicality	133.58	12.72	10.50	.<.001
Consistency*Lexicality	12.09	11.31	1.07	.285

Random effects	*Group*	*Variance*	*SD*	*Correlation*	
Item	Intercept	9.382e+03	96.86		
Participant	Intercept	3.541e+04	188.18		
	Consistency	3.392e+03	58.24	0.13	0.34
	Lexicality	3.531e+03	59.43	0.03	

Note: SE= standard error; Pr (>|t|) p-value calculated from the t-value; SD = standard deviation.

4.7.2 – Results of the picture naming task (children)

Compared to the RAT, the PNT was more effortful for the children as their accuracy rate was 79.44%. The lower accuracy might have been related to the task demand because children had to remember the label they used during the familiarization session, especially if the experimenter indicated the correct name of a picture when it was named incorrectly by them. Looking at the stimuli, three words (i.e., *duna-'dune'*, *duelo-'duel'* and *kilo*) were produced correctly below the chance level and therefore excluded from the analysis (5% of the collected data). Since it was not clear whether the incorrect productions made in the task were related to the OCE or to a wrong word choice, no statistical analysis was conducted to investigate this variable.

Concerning the SOT analysis, the data were trimmed like in the RAT resulting in the removal of 1.1% of the correct responses. Descriptively, the mean SOT of the consistent and inconsistent words did not seem to differ as it is shown in Table 14.

Table 14. Mean SOT of the PNT by consistency (SOT in milliseconds).

	Spanish words
Consistent	1024 (56)
Inconsistent	1023 (58)
Δ	-1

Note: Delta expresses the difference in ms between inconsistent and consistent items. SD in parentheses

The statistical analysis of the SOT data was also implemented with a GLMM in which the predicted variable was *SOT* (expressed in ms) and the predictor was *Consistency* (consistent = -1, inconsistent = 1) which was contrast coded. Random intercepts for *Participant* and *Stimulus* with a by-*Participant* random slope for *Consistency* were also included. The GLMM did not detect any main effects or interaction.

As in the PNT of Experiment 4, the PNT was further explored by running four separated GLMMs. Each model contained one of the co-variates based on the subjective ratings: *Imageability*, *Concreteness*, *Familiarity,* and *AoA*. The *performance* package in R (Lüdecke et al., 2021) revealed that the best-fitted model was the one including *Concreteness* as a co-variate. The model detected only a main effect for *Concreteness* but no main effect for *Consistency* and no interaction between the two variables.

4.8 – Summary of Experient 5

Experiment 5 investigated whether Spanish-speaking seven-year-olds (i.e., children at an early stage of reading acquisition) show an OCE in language production employing reading aloud and picture naming tasks. In contrast to adults, children rely more on the sub-lexical route during reading since they employ phonological decoding as a consequence of the stabilization of phoneme-to-grapheme mappings. Overall, the children showed a strong OCE in the RAT in terms of accuracy and SOT. Interestingly, there was no interaction between *Consistency* and *Lexicality* which implies that decoding was used regardless of whether the children had to read out loud a word or a pseudoword. This is striking also because the words employed in the RAT were high-frequency and they could have been processed through the lexical route (cf. Gerth & Festman, 2021). Also in this case, the present results fit current models of reading aloud like the DRC (Coltheart & Rastle, 1994).

Against my expectations, the children did not show any OCE in the PNT. This raises an important issue concerning the automaticity of the OCE in language production. This finding goes in the direction of Roelofs' argument (2006) that the OCE occurs only in those tasks in which orthography is relevant. Yet, as shown in studies comparing dyslexic and non-dyslexic children (Goswami et al., 1999; Swan & Goswami, 1997), orthography does play a role in PNT during phonological retrieval. An alternative explanation of why no OCE was found in the children's PNT could be that the children retrieved verbatim the labels of the pictures without drawing them from their lexicon but rather from their short-term memory. In fact, before the actual task, the children underwent a familiarization session in which they had to name all the pictures. This hypothesis suggests that the employed procedure represents a limitation of this study.

4.9 – General discussion

This chapter reports two experiments that investigated whether there is an OCE in language production. The issue was addressed from two perspectives: one cross-linguistic and the other related to the reading experience. Cross-linguistically, the pattern of results did not show any difference because neither French-speaking nor Spanish-speaking adults showed an OCE in both RAT and PNT. A numerical tendency (French Δ = 31, Spanish Δ = 32) in favor of an OCE in pseudoword reading aloud was found in both languages which suggests that probably both language groups employed similar processing strategies. The non-significance of the results, however, allows only for speculations. The fact that only high-frequency words were employed in both RAT and PNT might represent a limitation of the present study, but so were the stimuli of Experiment 1 which showed an OCE in auditory language perception. It should be noted, moreover, that the stimuli needed to be suitable also for seven-year-olds which implied the selection of quite frequent and common words. Future research should consequently explore whether the OCE is at least the result of strategic processing for the production of low-frequency words.

From the reading proficiency perspective, adults and children showed a different pattern in the RAT paradigm. While adults did not show any OCE in the task, children displayed a strong OCE. Yet, it is difficult to conclude that the OCE occurs in children's language processing in general since the effect was found only in the task in which orthography was relevant but not in the task in which orthography was not relevant (i.e., PNT). I restrict myself to pointing out that reading proficiency does make a difference in the way that speakers utilize the orthographic information related to the phonological representations in language production. Overall, the OCE can be seen as an effect of literacy on language production at least at the early stages of literacy acquisition and when orthography is present in the task.

Lastly, the present findings can be interpreted in light of current models of language production. I argue that models like the DRC (Coltheart et al., 2001b; Coltheart & Rastle, 1994) can explain the results of the RAT in both adults and children. As already mentioned, the numerical tendency found in consistent vs. inconsistent pseudowords in the RAT with adults and the strong OCE in both word and pseudoword production in children go in the direction of the bidirectional interactive processing at the sub-lexical level that the DRC model posits (Rastle, 2016). This means that when it comes to producing a pseudoword or when producing language as a child, there is a co-activation of all possible orthographic representations of an inconsistent phoneme, at least to some extent. This co-activation probably does not slow down the utterance of the pseudowords significantly in adults probably because adults are used to decoding efficiently. Concerning word production in adults, as I have mentioned before, the lexical-semantic route was probably preferred also considering that the stimuli were high-frequency. Nevertheless, children probably used the non-lexical-semantic route. Evidence for lexical processing of words in RAT comes also from the strong *Lexicality* effect found in both French and Spanish adults. Lexical processing is considered faster than sub-lexical processing in models like the DRC.

Again, I claim that the same arguments should hold for the PNT. The discrepancy between the present PNT results and those of Rastle and colleagues (2011) is likely due to an important difference in the paradigm: the present PNT had pictures depicting high-frequency, common words whereas in Rastle and colleagues the stimuli were novel words that had just been integrated into the mental lexicon. Even though the participants of that study learned novel words during a series of familiarization sessions, one cannot exclude that the OCE derived from the strategic use of orthography to facilitate the lemma retrieval of the item, especially considering that orthography was presented during the learning sessions.

The results presented in this chapter raise an important issue related to the perception-production link in language processing. I will discuss about this link in the next chapter.

References

Alario, F. X., Perre, L., Castel, C., & Ziegler, J. C. (2007). The role of orthography in speech production revisited. *Cognition*, *102*(3), 464–475. https://doi.org/10.1016/j.cognition.2006.02.002

Barca, L., Burani, C., & Arduino, L. S. (2002). Word naming times and psycholinguistic norms for Italian nouns. *Behavior Research Methods, Instruments, & Computers*, *34*(3), 424–434. https://doi.org/10.3758/BF03195471

Boersma, P., & Weenink, D. (2022). *Praat: Doing phonetics by computer* (6.2.12) [Computer software]. http://www.praat.org/

Borgwaldt, S. R., Bolger, P., & Jakab, E. (2010). Shallow Versus Deep Footprints in Pseudo-Word Grapheme-to-Phoneme Conversion: Dutch and English. *Journal of Germanic Linguistics*, *22*(4), 425–443. https://doi.org/10.1017/S1470542710000139

Coltheart, M., & Rastle, K. (1994). Serial processing in reading aloud: Evidence for dual-route models of reading. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(6), 1197–1211. https://doi.org/10.1037/0096-1523.20.6.1197

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A Dual Route Cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*, 204–256. https://doi.org/10.1037/0033-295X.108.1.204

Damian, M. F., & Bowers, J. S. (2003). Effects of orthography on speech production in a form-preparation paradigm. *Journal of Memory and Language*, *49*(1), 119–132. https://doi.org/10.1016/S0749-596X(03)00008-1

Duñabeitia, J. A., Crepaldi, D., Meyer, A. S., New, B., Pliatsikas, C., Smolka, E., & Brysbaert, M. (2018). MultiPic: A standardized set of 750 drawings with norms for six European languages. *Quarterly Journal of Experimental Psychology*, *71*(4), 808–816. https://doi.org/10.1080/17470218.2017.1310261

Gerth, S., & Festman, J. (2021). Reading Development, Word Length and Frequency Effects: An Eye-Tracking Study with Slow and Fast Readers. *Frontiers in Communication*, *6*, 1–12. https://doi.org/10.3389/fcomm.2021.743113

Goswami, U., Schneider, W., & Scheurich, B. (1999). Picture naming deficits in developmental dyslexia in German. *Developmental Science*, *2*(1), 53–58. https://doi.org/10.1111/1467-7687.00054

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., & Havinga, J. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review*, *98*(1), 122–142. https://doi.org/10.1037/0033-295X.98.1.122

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology*, *6*, 1171. https://doi.org/10.3389/fpsyg.2015.01171

Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., & Makowski, D. (2021). performance: An R Package for Assessment, Comparison and Testing of Statistical Models. *Journal of Open Source Software*, *6*(60), 3139. https://doi.org/10.21105/joss.03139

Meyer, A. S. (1990). The time course of phonological encoding in language production: The encoding of successive syllables of a word. *Journal of Memory and Language*, *29*(5), 524–545. https://doi.org/10.1016/0749-596X(90)90050-A

Peereman, R., & Content, A. (1999). LEXOP: A lexical database providing orthography-phonology statistics for French monosyllabic words. *Behavior Research Methods, Instruments, & Computers*, *31*(2), 376–379. https://doi.org/10.3758/BF03207735

Perret, C., & Bonin, P. (2019). Which variables should be controlled for to investigate picture naming in adults? A Bayesian meta-analysis. *Behavior Research Methods*, *51*(6), 2533–2545. https://doi.org/10.3758/s13428-018-1100-1

Powell, M. J. D. (2009). The BOBYQA algorithm for bound constrained optimization without derivatives. *Department of Applied Mathematics and Theoretical Physics, University of Cambridge (UK). Technical Report*, *NA2009/06*.

Qu, Q., & Damian, M. F. (2019). Orthographic effects in Mandarin spoken language production. *Memory & Cognition*, *47*(2), 326–334. https://doi.org/10.3758/s13421-018-0868-7

Rastle, K. (2016). Visual Word Recognition. In *Neurobiology of Language* (pp. 255–264). Elsevier. https://doi.org/10.1016/B978-0-12-407794-2.00021-3

Rastle, K., McCormick, S. F., Bayliss, L., & Davis, C. J. (2011). Orthography influences the perception and production of speech. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(6), 1588–1594. https://doi.org/10.1037/a0024833

Roelofs, A. (2006). The influence of spelling on phonological encoding in word reading, object naming, and word generation. *Psychonomic Bulletin & Review*, *13*(1), 33–37. https://doi.org/10.3758/BF03193809

RSTudio Team. (2020). *RStudio: Integrated Development for R* (1.3.1073) [Computer software]. RStudio, Inc. https://rstudio.com

Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, *94*(2), 143–174. https://doi.org/10.1348/000712603321661859

Swan, D., & Goswami, U. (1997). Picture naming deficits in developmental dyslexia: The phonological representations hypothesis. *Brain and Language*, *56*(3), 334–353. https://doi.org/10.1006/brln.1997.1855

Ziegler, J. C., & Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, *131*(1), 3–29. https://doi.org/10.1037/0033-2909.131.1.3

Chapter 5 – General discussion

The protagonist of the scientific narrative presented in this work is the Orthographic Consistency Effect (OCE). However, the OCE was used as a proxy to investigate the influence of literacy on auditory language perception and language production given that it is a by-product of literacy. To explore the OCE, I adopted two research approaches: cross-linguistic comparisons and developmental. Concerning the first, I contrasted French and Spanish because they differ in the opacity of their writing systems and they are consequently acquired through different processes (Seymour et al., 2003). As for the second approach, I tested children at an early stage of reading acquisition to

In this section, I will first summarize the findings of the five experiments presented in this dissertation, then I will connect the implications of the findings with each other and relate them to the theory. I will then draw some conclusions indicating possible future directions.

5.1. – Summary of the findings

In Experiment 1, I addressed the issue of whether the OCE in auditory language perception can be detected at the phonemic level in both French and Spanish. The results show that the OCE was present in French auditory pseudoword processing and Spanish auditory word processing. The cross-linguistic difference was interpreted in light of the degree of inconsistency in phoneme-to-grapheme mappings in the two languages. On the one hand, French, being an opaque language, requires more orthographic processing in auditory language perception because it helps identify the right lexical entry (Pattamadilok et al., 2009). On the other hand, Spanish as a transparent language seems to be less affected by orthographic inconsistencies but nevertheless shows an OCE in auditory word processing.

Experiment 2 explored the OCE at the phonemic level in Spanish-speaking seven-year-olds who were at an early stage of reading acquisition. The findings showed that children were more accurate at recognizing consistent than inconsistent words and pseudowords.

In Experiment 3, I addressed the issues of the timing and the underlying brain networks of the OCE by testing Spanish adults in an LDT and passive listening task while their brain activity was recorded using MEG. The first issue addressed using this technique was related to the pre-

/post-lexical debate (i.e., whether the OCE in auditory LDT occurs before or after lexical access). The second concern was to understand the brain correlates that underlie the OCE which, in turn, also raises the issue on the nature of phonemic representations (i.e., whether they are 'imbued' with orthographic information or whether they co-activate separated orthographic representations). The results showed that the OCE occurs at around 50-150ms post stimulus onset—as the signal was enhanced for inconsistent than consistent words in this time-window. As for the brain networks involved, a spatio-temporal sensor analysis showed that the cluster of MEG sensors involved was in the left fronto-temporal region, suggesting that phonemic representations are restructured with orthographic information as a consequence of literacy. To test the pervasiveness of the OCE in auditory language percpeiton, I ran a passive listening task that showed no OCE in either auditory word or pseudoword perception. The findings opened to possible interpretations on what kind of phonological representations that are involved based on the task.

Experiment 4 was dedicated to the exploration of the OCE in adult language production, with both French and Spanish as target languages, employing a reading aloud task (RAT) and picture naming task (PNT). The results showed only a numerical tendency for the OCE in pseudoword production in the RAT and no OCE in word production in either RAT or PNT. These findings are in line with previous literature (Alario et al., 2007; Roelofs, 2006).

Finally, Experiment 5 investigated the same psycholinguistic aspects as Experiment 4 but in Spanish-speaking seven-year-olds also using a RAT and a PNT. The findings revealed an OCE in both word and pseudoword production in the RAT but no OCE in word production in the PNT. Children seemed to rely on orthographic information related to phonemes, which facilitates lexical access in this stage of consolidation of orthographic representations

5.2 – Theoretical implications of the results for the auditory language perception literature

Experiments 1, 2, and 3 showed that the OCE is a robust psycholinguistic phenomenon in auditory language perception that also emerges at the phonemic level and is not limited to suprasegmental processing, as has previously been shown (Pattamadilok et al., 2007b; Ventura et al., 2007; Ziegler & Ferrand, 1998). I will address two theoretical issues that stem from these results: 1) The role (and the nature) of phonemes in auditory language perception and 2) the influence of literacy on speech in general (i.e., perception and production).

My findings raise the issue of what phonemes and graphemes represent in auditory language perception: Are they the most basic perceptual units in auditory language perception? As argued by Samuel (2020), there is an improper tendency in psycholinguistic research to consider abstract linguistic units as perceptual entities in the real world. Following his line of reasoning, one should not think that listeners perceive phonemes, syllables, or rimes. Rather, one should take these concepts as 'classifiers' that are employed in speech processing. The present findings show that orthographic information, in this case, graphemes, cooperate with phonemes in the abstraction and categorization of acoustic cues which is the fundamental part of auditory language processing. Interestingly, Morais (2021) argued that phonemes are a by-product of literacy. Many linguists define a *phoneme* as an abstract representation of speech sounds (i.e., *phones*) in a given language. According to Morais (2021), these representations are "forged and consolidated by literacy practice" (p. 5). As controversial as it sounds, my findings could be read in the light of his claim. Experiment 3 showed, even though preliminary, that phonemes are 'imbued' with orthographic information since the neural correlates of the OCE were located in the left fronto-temporal region. One could argue that this contamination underlies the more general phenomenon that phonemes are determined through reading acquisition. In other words, literacy acquisition can be seen as not only the establishment of a connection between sounds and spelling but also as the generation of abstract categories for those sounds. It is not my intention to address this 'chicken or egg' issue in much detail because it goes beyond the more modest scope of this work. Yet, it is insightful to look from Morais' angle what the present findings could mean for the general understanding of auditory language perception. Based on these considerations, one can be sure about one specific hypothesis: most likely, the pattern of results shown in this works would not be replicated with illiterate people.

A final issue that raises from Experiment 3 is related to whether there are multiple phonemic representations. The classification between speech-based and abstract phonemic representations (see Friedrich, 1990) seems to explain the difference found between the LDT and the passive listening task of Experiment 3. Again, an LDT requires an active (meta)linguistic processing that finalizes in lexical acces and, therefore, it could activate more abstract phonological representations that also contain orthographic representations. This type of phonological representations could be the actual result of literacy with respect to those which are more speech-based. Moreover, speech-based phonological representations could be at the base of low-level phonological processing underlying the passive listening task. This theoretical interpretation, however, raises an important issue: Why would children rely more on abstract phonological representations, given the results of Experiment 2 and 4, since they are still consolidating them? Also in this case, it could depend on the cognitive cost that the task requires. An LDT or a RAT require probably more cognitive effort to a seven-year-old than to an adult. The first task implicitly tests the vocabulary inventory of the child whereas the second implicitly tests their reading abilities. More research is needed to explore this theoretical interpretation especially from a developmental perspective.

My findings provide an important contribution to the understanding of how literacy influences auditory language perception. I argue that literacy does influence auditory language perception but not in a universal and univocal way. Its impact is modulated by the target language and the auditory input that the listener is processing. This means that literacy impacts auditory language perception to the extent that listeners utilize phoneme-to-grapheme mappings when they can help to optimize language processing. For example, in an LDT paradigm, French listeners rely on phoneme-to-grapheme mappings if they have to recognize a pseudoword because the phoneme-grapheme conversions will determine more efficiently if the stimulus is an actual word or not. Yet, they probably do not need to decode phoneme by phoneme if they have to recognize a word since they probably rely on bigger linguistic units (e.g., rimes). It should be clear, however, that this strategic utilization of phoneme-to-grapheme mappings is not conscious, but it simply derives from top-down processes occurring in auditory language perception. More concretely, this implies that whenever an auditory stimulus is perceived, its processing is determined not only by the incremental bottom-up flow in which the acoustic cues of speech are translated into abstract linguistic units but also that the lexical representations that are incrementally activated, in turn, contribute to the completion of the process.

From a modeling perspective, my findings seem to support interactive models like the bimodal interactive activation model (Grainger & Ferrand, 1996). In contrast to modular/autonomous models (e.g., Norris et al., 2000), interactive models posit top-down processes in auditory language perception. Taking the pattern of results of Experiments 1, 2 and 3, one can interpret both behavioral and, to some extent, the neuroimaging results as the evidence that both on the horizontal but also on the vertical level there is a flow of information that contributes to auditory (pseudo)word recognition. The horizontal flow of information would be between phoneme and graphemes while the vertical would be from phonological or orthographic word representations on the respective sub-lexical units.

Yet, these models do not explain what I argued before: the 'strategic' nature of the OCE based on the transparency of the writing system and based on the stimulus in analysis (i.e., word or pseudoword) and let alone the possible involvement of different types of phonological representations. In the case of Grainger and Ferrand's model (1996), one could argue that there might be an activation threshold which the different elements of the model have to reach to 'inform' each other. To my knowledge, there is strikingly no study that has ever systematically simulated with computational models the phonology-orthography interface in the auditory modality. This aspect should surely be addressed in future research.

The argument of the "strategic" nature of the OCE seems also to apply to the task requirements. The absence of the OCE in the passive listening task reported in Experiment 3 casts a shadow on the finding that orthographic information (i.e., literacy) pervasively affects auditory speech processing (cf. Pattamadilok et al., 2014). The task requirements of the LDT imply that participants know that they will hear something that does or does not match a lexical entry in their lexicon. Consequently, orthographic information is at play because it helps the recognition process. Conversely, during passive listening this information could be not necessary for auditory language perception. As already mentioned, bearing in mind that the passive listening task results are preliminary, another possible explanation of this pattern of results can be related to the language in question, namely Spanish. Further cross-linguistic research should systematically investigate this aspect.

5.3 – Theoretical implication of the results within the language production literature

Experiments 4 and 5 showed that literacy does not impact language production in the same way as it does auditory language perception. Only children seem to rely on orthographic information when producing language, but only when orthography is relevant in the task. This pattern of the results raises an important issue regarding the so-called "perception-production link". As introduced in Chapter 1, the perception-production link is a theoretical assumption according to which these two psycholinguistic domains are strictly connected. This assumption was translated in the present work into an attempt to investigate the OCE in both domains. The basic hypothesis was that if the OCE was detected in one domain, it should also be detected in the other. Yet, the findings of Experiment 3 suggest that the OCE emerges when the phonology-orthography interface facilitates the completion of a task. I argue that this explanation could be applied also to the findings of Experiment 4 and 5. No OCE was found in adults' language production tasks because phonology-to-orthography mappings were not necessarily activated during the production task or, more probably, they did not reach a level of activation that could impact reading aloud and picture naming. Considering interactive models of language productions, this pattern of results could be, in fact, interpreted in the light of the dual-route cascaded (DRC) model (Coltheart et al., 2001b; Coltheart & Rastle, 1994). Adults probably rely on the lexical-semantic route when producing language because it is considered quicker and more efficient. As for children, the clear OCE in the reading aloud task is probably more related to developmental reasons (i.e., pupils are more exposed and more used to decoding) than to the automaticity of the OCE per se. Nevertheless, the present work extends on the scarce knowledge of orthographic effects in language production by showing this difference between adults and children. As I have already argued, both the cross-linguistic and the developmental approach to the study of the OCE are the only systematic way to get a fuller understanding of the impact of literacy in language perception and production. In this case, the developmental approach has provided interesting and novel insights.

5.4 – Future directions

The present work investigated the role of literacy in auditory language perception and production by looking at the Orthographic Consistency Effect. I argued that the OCE can be seen as a proxy of the impact of literacy, yet it does not represent the complexity of literacy. Future research should systematically address the following issues:

- Compare the different linguistic levels (i.e., phonological, syllabic, etc.) in relation to the OCE both cross-linguistically and developmentally;
- Employ different experimental paradigms to investigate the OCE in both language perception and production;
- From a neuroimaging perspective, investigate the task-bias issue by observing neural-oscillations indicating top-down processes that modulate the impact of literacy on specific psycholinguistic domains.

5.5 – Conclusions

The present thesis confirms that the impact of literacy on linguistic domain like auditory perception and production exists but it is more complex than it has been depicted so far in the literature. On the one hand, the cross-linguistic approach employed in this work suggests that being literate in a language like French and in a language like Spanish can vary the way with which linguistic abstract units are organized and processed. On the other hand, the developmental approach confirms that becoming literate makes humans more efficient in the way they process language and deal with it in general. More specifically, it showed that phoneme-to-grapheme relations do not always contribute to the successful completion of a linguistic process. Finally, the combination of behavioral and neuroimaging methods also showed that the impact of literacy is complex and multifaceted. I argue that only by combining behavioral, neuroimaging and modelling approaches it will be possible to advance in understanding how this unique characteristic of humans, namely literacy, impacts the language system.

References

Alario, F. X., Perre, L., Castel, C., & Ziegler, J. C. (2007). The role of orthography in speech production revisited. *Cognition*, *102*(3), 464–475. https://doi.org/10.1016/j.cognition.2006.02.002

Coltheart, M., & Rastle, K. (1994). Serial processing in reading aloud: Evidence for dual-route models of reading. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(6), 1197–1211. https://doi.org/10.1037/0096-1523.20.6.1197

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A Dual Route Cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*, 204–256. https://doi.org/10.1037/0033-295X.108.1.204

Friedrich, F. J. (1990). Multiple phonological representations and verbal short-term memory. In G. Vallar & T. Shallice (Eds.), *Neuropsychological Impairments of Short-Term Memory* (pp. 74–93). Cambridge University Press. https://doi.org/10.1017/CBO9780511665547.005

Grainger, J., & Ferrand, L. (1996). Masked orthographic and phonological priming in visual word recognition and naming: Cross-task comparisons. *Journal of Memory and Language*, *35*(5), 623–647. https://doi.org/10.1006/jmla.1996.0033

Morais, J. (2021). The phoneme: A conceptual heritage from alphabetic literacy. *Cognition*, *213*, 104740. https://doi.org/10.1016/j.cognition.2021.104740

Norris, D., McQueen, J. M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences*, *23*(3), 299–325. https://doi.org/10.1017/S0140525X00003241

Pattamadilok, C., Morais, J., Colin, C., & Kolinsky, R. (2014). Unattentive speech processing is influenced by orthographic knowledge: Evidence from mismatch negativity. *Brain and Language*, *137*, 103–111. https://doi.org/10.1016/j.bandl.2014.08.005

Pattamadilok, C., Morais, J., De Vylder, O., Ventura, P., & Kolinsky, R. (2009). The orthographic consistency effect in the recognition of French spoken words: An early developmental shift from sublexical to lexical orthographic activation. *Applied Psycholinguistics*, *30*(3), 441–462. https://doi.org/10.1017/S0142716409090225

Pattamadilok, C., Morais, J., Ventura, P., & Kolinsky, R. (2007). The locus of the orthographic consistency effect in auditory word recognition: Further evidence from French. *Language and Cognitive Processes*, *22*(5), 700–726. https://doi.org/10.1080/01690960601049628

Roelofs, A. (2006). The influence of spelling on phonological encoding in word reading, object naming, and word generation. *Psychonomic Bulletin & Review*, *13*(1), 33–37. https://doi.org/10.3758/BF03193809

Samuel, A. G. (2020). Psycholinguists should resist the allure of linguistic units as perceptual units. *Journal of Memory and Language*, *111*, 104070. https://doi.org/10.1016/j.jml.2019.104070

Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, *94*(2), 143–174. https://doi.org/10.1348/000712603321661859

Ventura, P., Morais, J., & Kolinsky, R. (2007). The development of the orthographic consistency effect in speech recognition: From sublexical to lexical involvement. *Cognition*, *105*(3), 547–576. https://doi.org/10.1016/j.cognition.2006.12.005

Ziegler, J. C., & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin and Review*, *5*(4), 683–689. https://doi.org/10.3758/BF03208845

Appendix A – Word matching within language and across conditions (consistent/inconsistent) in Experiment 1 and 2 (LDT)

French word stimuli matching – TOST Results

		t	df	p
W_frq (Zipf)	t-test	0.2856	58.0	0.776
	TOST Upper	-2.565	58.0	0.006
	TOST Lower	3.14	58.0	0.001
Num. letters	t-test	-0.3681	57.0	0.714
	TOST Upper	-3.219	57.0	0.001
	TOST Lower	2.48	57.0	0.008
Orth. neigh.	t-test	1.4681	51.5	0.148
	TOST Upper	-1.382	51.5	0.086
	TOST Lower	4.32	51.5	< .001
Orth. uniqueness point	t-test	-0.5757	50.8	0.567
	TOST Upper	-3.426	50.8	< .001
	TOST Lower	2.27	50.8	0.014
Num. phonemes	t-test	-0.7538	52.9	0.454
	TOST Upper	-3.604	52.9	< .001
	TOST Lower	2.10	52.9	0.020
Num. syllables	t-test	-1.0947	54.4	0.278
	TOST Upper	-3.945	54.4	< .001
	TOST Lower	1.76	54.4	0.042
Num. homophones	t-test	0.0653	50.0	0.948
	TOST Upper	-2.785	50.0	0.004
	TOST Lower	2.92	50.0	0.003

Spanish word matching – TOST Results

		t	df	p
W_frq (Zipf)	t-test	1.4651	56.5	0.148
	TOST Upper	-1.39	56.5	0.086
	TOST Lower	4.32	56.5	< .001
Num. letters	t-test	0.2398	56.6	0.811
	TOST Upper	-2.61	56.6	0.006
	TOST Lower	3.09	56.6	0.002
Orth. neigh.	t-test	0.6560	56.9	0.514
	TOST Upper	-2.19	56.9	0.016
	TOST Lower	3.51	56.9	< .001
Orth. uniqueness point	t-test	1.3023	57.0	0.198
	TOST Upper	-1.55	57.0	0.064
	TOST Lower	4.15	57.0	< .001
Num. phonemes	t-test	0.7276	55.5	0.470
	TOST Upper	-2.12	55.5	0.019
	TOST Lower	3.58	55.5	< .001
Num. syllables	t-test	1.1682	57.9	0.248
	TOST Upper	-1.68	57.9	0.049
	TOST Lower	4.02	57.9	< .001
Num. homophones	t-test	-0.7504	55.5	0.456
	TOST Upper	-3.60	55.5	< .001
	TOST Lower	2.10	55.5	0.020

Appendix A – Word matching within language and across conditions (consistent/inconsistent) in Experiment 1 and 2 (LDT)

Phon. neigh.	t-test	0.0689	56.9	0.945
	TOST Upper	-2.782	56.9	0.004
	TOST Lower	2.92	56.9	0.003
Phon. uniqueness point	t-test	-1.1409	49.2	0.259
	TOST Upper	-3.991	49.2	< .001
	TOST Lower	1.71	49.2	0.047
Mean biphone frq	t-test	-1.5201	54.7	0.134
	TOST Upper	-4.371	54.7	< .001
	TOST Lower	1.33	54.7	0.094
Mean audio duration	t-test	0.0831	55.7	0.934
	TOST Upper	-2.767	55.7	0.004
	TOST Lower	2.93	55.7	0.002

Phon. neigh.	t-test	-0.2216	57.9	0.825
	TOST Upper	-3.07	57.9	0.002
	TOST Lower	2.63	57.9	0.005
Phon. uniqueness point	t-test	1.5496	56.3	0.127
	TOST Upper	-1.30	56.3	0.099
	TOST Lower	4.40	56.3	< .001
Mean biphone frq.	t-test	-0.0477	57.4	0.962
	TOST Upper	-2.90	57.4	0.003
	TOST Lower	2.80	57.4	0.003
Mean audio duration	t-test	-0.7326	56.7	0.467
	TOST Upper	-3.58	56.7	< .001
	TOST Lower	2.12	56.7	0.019

Appendix A – Pseudoword matching within language and across conditions (consistent/inconsistent) in Experiment 1 and 2 (LDT)

French pseudoword matching -TOST Results

		t	df	p
Mean biphone frq	t-test	-0.0652	55.1	0.948
	TOST Upper	-2.92	55.1	0.003
	TOST Lower	2.79	55.1	0.004
Mean phoneme freq	t-test	-1.4345	55.8	0.157
	TOST Upper	-4.29	55.8	< .001
	TOST Lower	1.42	55.8	0.081
Num. phonemes	t-test	-0.6146	54.7	0.541
	TOST Upper	-3.47	54.7	< .001
	TOST Lower	2.24	54.7	0.015
Phon. neighbors (count)	t-test	-1.1020	54.9	0.275
	TOST Upper	-3.95	54.9	< .001
	TOST Lower	1.75	54.9	0.043
Mean neighbor frq	t-test	1.2609	29.2	0.217
	TOST Upper	-1.59	29.2	0.061
	TOST Lower	4.11	29.2	< .001
Mean audio duration	t-test	-1.5110	58.0	0.136
	TOST Upper	-4.36	58.0	< .001
	TOST Lower	1.34	58.0	0.093

Spanish pseudoword matching -TOST Results

		t	df	p
Mean biphone frq.	t-test	-0.513	41.5	0.611
	TOST Upper	-3.36	41.5	< .001
	TOST Lower	2.34	41.5	0.012
Mean phoneme frq.	t-test	1.310	48.6	0.196
	TOST Upper	-1.54	48.6	0.065
	TOST Lower	4.16	48.6	< .001
Num. phonemes	t-test	0.728	55.5	0.470
	TOST Upper	-2.12	55.5	0.019
	TOST Lower	3.58	55.5	< .001
Phon. neighbors (count)	t-test	-0.301	55.0	0.765
	TOST Upper	-3.12	55.0	0.001
	TOST Lower	2.52	55.0	0.007
Mean neighbor frq.	t-test	-0.718	32.1	0.478
	TOST Upper	-3.52	32.1	< .001
	TOST Lower	2.09	32.1	0.022
Mean audio duration	t-test	0.758	57.6	0.452
	TOST Upper	-2.09	57.6	0.020
	TOST Lower	3.61	57.6	< .001

Note: all t-test are Welch's t-tests. The stimuli matching was checked with the TOSTER package (Lakens, 2016) in Jamovi (The Jamovi Project, 2021). The Upper and Lower limits of the TOST test were determined by means of a predetermined effect size calculated with the jpower package (Morey & Selker, 2021). If the p-value of both limits is <.05, the assumption of equality is accepted. In some cases, one of the two limits is not statistically significant. In that case, if the p-value of the Welch's t-test is >.05, it was assumed that the stimuli were matched on that variable.

Appendix B – Stimuli matching across languages (French/Spanish) in Experiment 1

French-Spanish word matching - One-Way ANOVA (Welch's)

	F	df1	df2	p
W_frq (Zipf)	1.168	3	64.1	**0.329**
Num. letters	0.506	3	63.8	**0.680**
Orth. Neigh.	6.295	3	62.6	< .001
Orth. uniqueness point	7.820	3	63.9	< .001
Num. phonemes	8.025	3	63.9	< .001
Num. syllables	16.059	3	63.9	< .001
Num. homophones	45.292	3	58.3	< .001
Phon. Neigh.	3.086	3	63.8	0.033
Phon. uniqueness point	38.776	3	63.7	< .001
Mean biphone frq.	9.560	3	60.1	< .001
Mean audio duration	0.338	3	64.0	**0.798**

Note: *p*-values in bold indicated matched variable

French-Spanish pseudoword matching - One-Way ANOVA (Welch's)

	F	df1	df2	p
Mean biphone frq.	4.14	3	62.9	0.010
Mean phoneme frq.	18.92	3	63.6	< .001
Num. phonemes	4.43	3	63.9	0.007
Phon. neighbors (count)	7.65	3	59.1	< .001
Mean neighbor frq.	1.19	3	48.1	**0.322**
Mean audio duration	3.12	3	64.4	0.032

Note: *p*-values in bold indicated matched variable

Note: Cross-linguistic stimuli matching tested with Welch's ANOVA across conditions (i.e., French consistent, French inconsistent, Spanish consistent, Spanish inconsistent).

Appendix C – Word matching across conditions (consistent/inconsistent) in Experiment 3 (MEG)

Spanish Word Matching MEG - TOST Results

		t	df	p
Spanish/Basque cognates	t-test	1.8424	311	0.066
	TOST Upper	-2.147	311	0.016
	TOST Lower	5.83	311	< .001
Log word count	t-test	1.7109	302	0.088
	TOST Upper	-2.278	302	0.012
	TOST Lower	5.70	302	< .001
Word frq. per million	t-test	-0.4283	294	0.669
	TOST Upper	-4.417	294	< .001
	TOST Lower	3.56	294	< .001
Num. letters	t-test	-2.5435	316	0.011
	TOST Upper	-6.533	316	< .001
	TOST Lower	1.45	316	0.075
Orth. Neigh.	t-test	3.2282	297	0.001
	TOST Upper	-0.761	297	0.224
	TOST Lower	7.22	297	< .001
Higher Frq. Orth. Neigh.	t-test	1.1698	301	0.243
	TOST Upper	-2.819	301	0.003
	TOST Lower	5.16	301	< .001
Orth. uniqueness point	t-test	-0.7808	313	0.436
	TOST Upper	-4.770	313	< .001
	TOST Lower	3.21	313	< .001
Levensthein's distance	t-test	-2.7617	318	0.006
	TOST Upper	-6.751	318	< .001
	TOST Lower	1.23	318	0.110
Mean bigram frq. (token-positional)	t-test	0.4927	316	0.623
	TOST Upper	-3.496	316	< .001
	TOST Lower	4.48	316	< .001
Mean bigram frq. (type-positional)	t-test	-0.2437	299	0.808
	TOST Upper	-4.233	299	< .001
	TOST Lower	3.75	299	< .001
Num. phonemes	t-test	-1.5260	318	0.128
	TOST Upper	-5.515	318	< .001
	TOST Lower	2.46	318	0.007

Appendix C – Word matching across conditions (consistent/inconsistent) in Experiment 3 (MEG)

Spanish Word Matching MEG - TOST Results

		t	df	p
Accented syllable	t-test	-0.9947	316	0.321
	TOST Upper	-4.984	316	< .001
	TOST Lower	2.99	316	0.001
Num. Homophones	t-test	-1.4607	307	0.145
	TOST Upper	-5.450	307	< .001
	TOST Lower	2.53	307	0.006
Phon. Neigh.	t-test	1.1696	318	0.243
	TOST Upper	-2.819	318	0.003
	TOST Lower	5.16	318	< .001
Higher Frq. Phon. Neigh.	t-test	-0.1291	318	0.897
	TOST Upper	-4.118	318	< .001
	TOST Lower	3.86	318	< .001
Phon. uniqueness point	t-test	-1.1221	318	0.263
	TOST Upper	-5.111	318	< .001
	TOST Lower	2.87	318	0.002
Mean biphone frq. (token-positional)	t-test	1.7630	312	0.079
	TOST Upper	-2.226	312	0.013
	TOST Lower	5.75	312	< .001
Mean biphone frq. (type-positional)	t-test	-1.3487	293	0.178
	TOST Upper	-5.338	293	< .001
	TOST Lower	2.64	293	0.004
Audio duration	t-test	1.8057	137	0.073
	TOST Upper	-1.015	137	0.156
	TOST Lower	4.63	137	< .001

Note. Welch's t-test

Appendix C – Pseudoword matching across conditions (consistent/inconsistent) in Experiment 3 (MEG)

Spanish Pseudoword Matching MEG - TOST Results

		t	df	p
Mean bigram frq.	t-test	0.9647	306	0.335
	TOST Upper	-3.07	306	0.001
	TOST Lower	5.00	306	< .001
Mean gram frq.	t-test	2.7406	319	0.006
	TOST Upper	-1.29	319	0.099
	TOST Lower	6.77	319	< .001
Num. letters	t-test	-1.2711	324	0.205
	TOST Upper	-5.30	324	< .001
	TOST Lower	2.76	324	0.003
Orth. Neigh.	t-test	0.7392	324	0.460
	TOST Upper	-3.29	324	< .001
	TOST Lower	4.77	324	< .001
Frq. Orth. Neigh	t-test	1.3957	280	0.164
	TOST Upper	-2.62	280	0.005
	TOST Lower	5.41	280	< .001
Mean biphone frq.	t-test	-1.9083	270	0.057
	TOST Upper	-5.95	270	< .001
	TOST Lower	2.14	270	0.017
Mean phoneme frq.	t-test	0.6231	320	0.534
	TOST Upper	-3.41	320	< .001
	TOST Lower	4.65	320	< .001
Num. phonemes	t-test	-0.1329	324	0.894
	TOST Upper	-4.16	324	< .001
	TOST Lower	3.90	324	< .001
Phon. Neigh.	t-test	-1.1781	315	0.240
	TOST Upper	-5.21	315	< .001
	TOST Lower	2.86	315	0.002
Frq. Phon. Neigh.	t-test	-0.0973	310	0.923
	TOST Upper	-4.12	310	< .001
	TOST Lower	3.92	310	< .001

Note. Welch's t-test

Appendix C – Pseudoword matching across conditions (consistent/inconsistent) in Experiment 3 (MEG)

Appendix D – Word matching within languages and across conditions (consistent/inconsistent) in Experiment 5 and 6 (RAT)

French Word Matching (RAT) - TOST Results

		t	df	p
W_frq (Zipf)	t-test	0.1890	38.0	0.851
	TOST Upper	-2.69	38.0	0.005
	TOST Lower	3.06	38.0	0.002
Num. letters	t-test	-1.2764	38.0	0.210
	TOST Upper	-4.15	38.0	< .001
	TOST Lower	1.60	38.0	0.059
Orth. neigh.	t-test	-0.2530	38.0	0.802
	TOST Upper	-3.13	38.0	0.002
	TOST Lower	2.62	38.0	0.006
Orth. uniqueness point	t-test	-0.4704	38.0	0.641
	TOST Upper	-3.34	38.0	< .001
	TOST Lower	2.40	38.0	0.011
Num. phonemes	t-test	-1.3529	38.0	0.184
	TOST Upper	-4.23	38.0	< .001
	TOST Lower	1.52	38.0	0.068
Num. syllables	t-test	-0.7415	38.0	0.463
	TOST Upper	-3.62	38.0	< .001
	TOST Lower	2.13	38.0	0.020
Num. homophones	t-test	0.4078	37.0	0.686
	TOST Upper	-2.43	37.0	0.010

Spanish Matching Words (RAT) - TOST Results

		t	df	p
W_frq (Zipf)	t-test	0.562	38.0	0.577
	TOST Upper	-2.31	38.0	0.013
	TOST Lower	3.44	38.0	< .001
Num. letters	t-test	-0.340	38.0	0.735
	TOST Upper	-3.21	38.0	0.001
	TOST Lower	2.53	38.0	0.008
Orth. neigh.	t-test	0.551	38.0	0.585
	TOST Upper	-2.32	38.0	0.013
	TOST Lower	3.43	38.0	< .001
Orth. uniqueness point	t-test	0.317	38.0	0.753
	TOST Upper	-2.56	38.0	0.007
	TOST Lower	3.19	38.0	0.001
Num. phonemes	t-test	0.000	38.0	1.000
	TOST Upper	-2.87	38.0	0.003
	TOST Lower	2.87	38.0	0.003
Num. syllables	t-test	-0.876	38.0	0.387
	TOST Upper	-3.75	38.0	< .001
	TOST Lower	2.00	38.0	0.026
Num. homophones	t-test	-1.292	38.0	0.204
	TOST Upper	-4.17	38.0	< .001
	TOST Lower	1.58	38.0	0.061

Appendix D – Word matching within languages and across conditions (consistent/inconsistent) in Experiment 5 and 6 (RAT)

French Word Matching (RAT) - TOST Results

		t	df	p
	TOST Lower	2.96	38.0	0.003
Phon. uniqueness point	t-test	-0.9002	38.0	0.374
	TOST Upper	-3.77	38.0	< .001
	TOST Lower	1.97	38.0	0.028
Mean biphone frq.	t-test	-0.3015	38.0	0.765
	TOST Upper	-3.18	38.0	0.001
	TOST Lower	2.57	38.0	0.007
Mean bigram frq.	t-test	1.4058	38.0	0.168
	TOST Upper	-1.47	38.0	0.075
	TOST Lower	4.28	38.0	< .001

Spanish Matching Words (RAT) - TOST Results

		t	df	p
Phon. uniqueness point	t-test	0.000	38.0	1.000
	TOST Upper	-2.87	38.0	0.003
	TOST Lower	2.87	38.0	0.003
Mean biphone frq.	t-test	1.754	38.0	0.087
	TOST Upper	-1.12	38.0	0.135
	TOST Lower	4.63	38.0	< .001
Mean bigram frq.	t-test	1.070	37.0	0.291
	TOST Upper	-1.77	37.0	0.043
	TOST Lower	3.91	37.0	< .001

Appendix D – Word matching within languages and across conditions (consistent/inconsistent) in Experiment 5 and 6 (RAT)

French Matching Pseudowords (RAT) - TOST Results

		t	df	p
Mean biphone frq	t-test	-0.417	37.0	0.679
	TOST Upper	-3.25	37.0	0.001
	TOST Lower	2.421	37.0	0.010
Mean phoneme frq	t-test	-0.757	37.0	0.454
	TOST Upper	-3.59	37.0	< .001
	TOST Lower	2.080	37.0	0.022
Num. phonemes	t-test	-1.873	37.0	0.069
	TOST Upper	-4.71	37.0	< .001
	TOST Lower	0.964	37.0	0.171
Phon. neighbors (count)	t-test	1.634	37.0	0.111
	TOST Upper	-1.20	37.0	0.118
	TOST Lower	4.471	37.0	< .001
Mean Phon. Neigh. frq.	t-test	0.373	37.0	0.711
	TOST Upper	-2.46	37.0	0.009
	TOST Lower	3.211	37.0	0.001
Mean bigram frq.	t-test	0.891	37.0	0.379
	TOST Upper	-1.95	37.0	0.030
	TOST Lower	3.728	37.0	< .001
Mean gram frq.	t-test	-0.280	37.0	0.781
	TOST Upper	-3.12	37.0	0.002

Spanish Matching Pseudowords (RAT) - TOST Results

		t	df	p
Mean biphone frq	t-test	-0.7309	38.0	0.469
	TOST Upper	-3.61	38.0	< .001
	TOST Lower	2.14	38.0	0.019
Mean phoneme frq	t-test	-0.3538	38.0	0.725
	TOST Upper	-3.23	38.0	0.001
	TOST Lower	2.52	38.0	0.008
Num. phonemes	t-test	-0.1184	38.0	0.906
	TOST Upper	-2.99	38.0	0.002
	TOST Lower	2.76	38.0	0.004
Phon. neighbors (count)	t-test	-0.4070	38.0	0.686
	TOST Upper	-3.28	38.0	0.001
	TOST Lower	2.47	38.0	0.009
Mean Phon. Neigh. frq.	t-test	-0.3406	38.0	0.735
	TOST Upper	-3.22	38.0	0.001
	TOST Lower	2.53	38.0	0.008
Mean bigram frq.	t-test	-0.0804	38.0	0.936
	TOST Upper	-2.95	38.0	0.003
	TOST Lower	2.79	38.0	0.004
Mean gram frq.	t-test	0.7361	38.0	0.466
	TOST Upper	-2.14	38.0	0.019
	TOST Lower	3.61	38.0	< .001
Num. letters	t-test	-0.3506	38.0	0.728
	TOST Upper	-3.23	38.0	0.001

Appendix D – Word matching within languages and across conditions (consistent/inconsistent) in Experiment 5 and 6 (RAT)

French Matching Pseudowords (RAT) - TOST Results

		t	df	p
	TOST Lower	2.558	37.0	0.007
Num. letters	t-test	-0.769	37.0	0.447
	TOST Upper	-3.61	37.0	< .001
	TOST Lower	2.068	37.0	0.023
Ortho. neighbors (count)	t-test	0.200	37.0	0.843
	TOST Upper	-2.64	37.0	0.006
	TOST Lower	3.037	37.0	0.002
Mean Ortho. Neigh. frq.	t-test	-0.593	36.0	0.557
	TOST Upper	-3.39	36.0	< .001
	TOST Lower	2.204	36.0	0.017

Spanish Matching Pseudowords (RAT) - TOST Results

		t	df	p
	TOST Lower	2.52	38.0	0.008
Ortho. neighbors (count)	t-test	-0.5941	38.0	0.556
	TOST Upper	-3.47	38.0	< .001
	TOST Lower	2.28	38.0	0.014
Mean Ortho. Neigh. frq.	t-test	-0.1110	38.0	0.912
	TOST Upper	-2.99	38.0	0.002
	TOST Lower	2.76	38.0	0.004

Appendix E – Word matching across languages in Experiment 4 (RAT)

French-Spanish Matching Words (RAT) - One-Way ANOVA (Welch's)

	F	df1	df2	p
Word frq. (Zipf)	0.154	1	74.8	**0.696**
Num. letters	0.199	1	76.0	**0.657**
Num. phonemes	16.869	1	77.1	< .001
Ortho. neighbors	8.694	1	63.0	0.004
Ortho. uniqueness point	11.416	1	77.9	0.001
Num. syllables	0.122	1	76.8	**0.728**
Homophones	0.508	1	77.2	**0.478**
Phono. neighbors	5.694	1	68.4	0.020
Phono. uniqueness point	14.199	1	75.7	< .001

French-Spanish Matching Pseudowords (RAT) - One-Way ANOVA (Welch's)

	F	df1	df2	p
Biphone frq.	11.43490	1	59.0	0.001
Phoneme frq.	34.15238	1	75.2	< .001
Num. phonemes	15.08850	1	76.4	< .001
Phono. neighbors	12.20907	1	47.8	0.001
Frq. phono neigh.	0.14994	1	73.7	**0.700**
Bigram frq.	1.09213	1	66.2	**0.300**
Gram frq.	0.00847	1	75.3	**0.927**
Num. letters	0.67068	1	69.4	**0.416**
Ortho. neigh.	0.69726	1	63.8	**0.407**
Frq. Ortho. neigh.	0.44372	1	56.0	**0.508**

Note: *p*-values in bold indicate matched variable.

Appendix G – Word matching within languages and across conditions (consistent/inconsistent) in Experiment 4 and 5 (PNT)

French Word Matching (PNT) - TOST Results

		t	df	p
Word frq (Zipf)	t-test	0.7299	58.0	0.468
	TOST Upper	-2.12	58.0	0.019
	TOST Lower	3.58	58.0	< .001
Num. letters	t-test	0.4339	58.0	0.666
	TOST Upper	-2.42	58.0	0.009
	TOST Lower	3.28	58.0	< .001
Ortho. neighbors	t-test	1.5499	58.0	0.127
	TOST Upper	-1.30	58.0	0.099
	TOST Lower	4.40	58.0	< .001
Ortho. uniqueness point	t-test	1.0104	58.0	0.316
	TOST Upper	-1.84	58.0	0.035
	TOST Lower	3.86	58.0	< .001
Num. phonemes	t-test	0.8646	58.0	0.391
	TOST Upper	-1.99	58.0	0.026
	TOST Lower	3.72	58.0	< .001
Num. syllables	t-test	-0.2266	58.0	0.822
	TOST Upper	-3.08	58.0	0.002
	TOST Lower	2.62	58.0	0.006
Num homophones	t-test	0.2612	58.0	0.795
	TOST Upper	-2.59	58.0	0.006
	TOST Lower	3.11	58.0	0.001
Phono. neighbors	t-test	-0.0626	58.0	0.950
	TOST Upper	-2.91	58.0	0.003

Spanish Word Matching (PNT) - TOST Results

		t	df	p
Word frq. (Zipf)	t-test	0.208	58.0	0.836
	TOST Upper	-2.64	58.0	0.005
	TOST Lower	3.06	58.0	0.002
Num. letters	t-test	-0.126	58.0	0.900
	TOST Upper	-2.98	58.0	0.002
	TOST Lower	2.72	58.0	0.004
Ortho. neighbors	t-test	-0.531	58.0	0.597
	TOST Upper	-3.38	58.0	< .001
	TOST Lower	2.32	58.0	0.012
Ortho. uniqueness point	t-test	-0.250	58.0	0.803
	TOST Upper	-3.10	58.0	0.001
	TOST Lower	2.60	58.0	0.006
Num. phonemes	t-test	-0.523	58.0	0.603
	TOST Upper	-3.37	58.0	< .001
	TOST Lower	2.33	58.0	0.012
Num. syllables	t-test	-0.317	58.0	0.753
	TOST Upper	-3.17	58.0	0.001
	TOST Lower	2.53	58.0	0.007
Num. homophones	t-test	1.240	58.0	0.220
	TOST Upper	-1.61	58.0	0.056
	TOST Lower	4.09	58.0	< .001

Appendix G – Word matching within languages and across conditions (consistent/inconsistent) in Experiment 4 and 5 (PNT)

French Word Matching (PNT) - TOST Results

		t	df	p
	TOST Lower	2.79	58.0	0.004
Phono. uniqueness point	t-test	0.2763	58.0	0.783
	TOST Upper	-2.57	58.0	0.006
	TOST Lower	3.13	58.0	0.001
Biphone frq	t-test	0.9953	52.0	0.324
	TOST Upper	-1.71	52.0	0.047
	TOST Lower	3.70	52.0	< .001
Visual complexity (MultiPic)	t-test	-0.9636	43.0	0.341
	TOST Upper	-3.40	43.0	< .001
	TOST Lower	1.48	43.0	0.074

Spanish Word Matching (PNT) - TOST Results

		t	df	p
Phono. neighbors	t-test	0.543	58.0	0.589
	TOST Upper	-2.31	58.0	0.012
	TOST Lower	3.39	58.0	< .001
Phono. uniqueness point	t-test	-0.523	58.0	0.603
	TOST Upper	-3.37	58.0	< .001
	TOST Lower	2.33	58.0	0.012
Bigram frq.	t-test	-0.997	58.0	0.323
	TOST Upper	-3.85	58.0	< .001
	TOST Lower	1.85	58.0	0.034
Visual complexity (Multipic)	t-test	1.468	51.0	0.148
	TOST Upper	-1.21	51.0	0.116
	TOST Lower	4.15	51.0	< .001

Appendix G – Word matching across languages (French/Spanish) in Experiment 4 (PNT)

One-Way ANOVA (Welch's)

	F	**df1**	**df2**	**p**
Word frq. (Zipf)	1.35e-5	1	115.3	**0.997**
Num. letters	3.62792	1	115.4	**0.059**
Num. phonemes	22.84369	1	117.6	< .001
Ortho. neighbors	25.65588	1	81.7	< .001
Ortho. uniqueness point	2.04818	1	105.4	**0.155**
Num. syllables	60.86316	1	106.9	< .001
Num. homophones	6.77838	1	115.2	0.010
Phono. neighbors	25.57041	1	81.4	< .001
Phono. uniqueness point	131.76179	1	117.6	< .001
Visual complexity (MultiPic)	0.00676	1	79.4	**0.935**

Appendix H – Stimuli lists of Experiment 3 (MEG)

French – Words		**Spanish – Words**		**French – Pseudowords**		**Spanish – Pseudowords**	
Consistent	*Inconsistent*	*Consistent*	*Inconsistent*	*Consistent*	*Inconsistent*	*Consistent*	*Inconsistent*
bâton	fils	daño	balcón	/damɔ̃/	/ʒiʒ/	/pafo/	/kalbon/
boule	flaque	dedal	billete	/duɲ/	/slas/	/pemal/	/kiʎeme/
boîte	forêt	defensa	bomba	/dwaʃ/	/ʒolɛ/	/pelensa/	/komka/
bonbon	pharmacie	deporte	bota	/dɔ̃mɔ̃/	/saʁnasi/	/peɲorte/	/kona/
bouchon	phare	deseo	brazo	/dulɔ̃/	/ʒal/	/peɾeo/	/kɾaxo/
bol	cousin	desorden	varita	/dɔʃ/	/sufɛ̃/	/pesorlen/	/kaɾifa/
brûlure	caisse	dolor	ventana	/tʁymyʁ/	/sɛf/	/ponoɾ/	/kenfana/
date	caméra	doña	verja	/mab/	/samena/	/pofa/	/keɾθa/
don	côte	duda	vida	/bə/	/soʃ/	/pusa/	/kifa/
doudoune	castor	dueño	voz	/bulun/	/saktɔʁ/	/pwemo/	/kox/
douleur	képi	palmera	caries	/bulœv/	/seli/	/dalneɾa/	/bafjes/
doute	quatre	pastor	carne	/bun/	/sadʁ/	/damtoɾ/	/baɾfe/
drame	kilo	pelea	corbata	/bʁan/	/ʒivo/	/demea/	/boɾxata/
lardon	quai	plancha	cobra	/paʁnɔ̃/	/zɛ̃/	/dwantʃa/	/boθɾa/
larme	question	planeta	corcho	/paʁd/	/sɛʒtjɔ̃/	/dɾaneta/	/boɾmo/
mâchoire	cerf	planta	kayak	/paʃwal/	/kɛʃ/	/dwanta/	/baxak/
mouche	cerveau	pluma	ketchup	/pub/	/kɛʁbo/	/dɾuma/	/beðlup/
morve	cible	potro	queja	/pɔʁl/	/kibʁ/	/dotno/	/bema/
madame	cigogne	pulsera	quema	/patam/	/kizɔɲ/	/dulfeɾa/	/bepa/
parole	cicatrice	puma	quinta	/nabɔl/	/kiʒatʁis/	/duta/	/binfa/

Appendix H – Stimuli lists of Experiment 3 (MEG)

peur	soupe	fecha	robot	/lœt/	/kum/	/nesa/	/θokot/
pneu	sabot	frase	rueda	/tʁø/	/fato/	/nwase/	/θwena/
peluche	sapin	fruta	cerdo	/məʁyʃ/	/kaʁɛ̃/	/luɾsa/	/xeɾlo/
porte	serrure	leche	cielo	/lɔʁv/	/kemyʁ/	/meɲe/	/rjemo/
preuve	sirène	leña	circo	/vʁœn/	/kiʁɛl/	/fela/	/xiɾbo/
puma	gifle	letra	zona	/lyva/	/sisl/	/mefɾa/	/xola/
roulade	gendarme	mantel	zumo	/bunad/	/fɑ̃daʁt/	/fansel/	/xuso/
ruche	girafe	mapa	jarra	/myv/	/kiʁag/	/nala/	/θaxa/
vapeur	jupe	mono	jarabe	/ʁanœʁ/	/fyd/	/lofo/	/θarake/
retour	jour	norte	jueves	/bətut/	/fub/	/foste/	/rwekes/

Appendix I – Stimuli lists of Experiment 3 (MEG)

Words		**Pseudowords**	
Consistent	*Inconsistent*	*Consistent*	*Inconsistent*
dado	banco	/dalon/	/bamia/
dama	baño	/daron/	/bedis/
diana	barba	/dasuar/	/bemblo/
diésel	barra	/dene/	/benlion/
dieta	bastón	/doda/	/bido/
dorso	beca	/dosa /	/bifo/
dote	bestia	/duenes/	/binla/
drama	bici	/dueso/	/bisan/
duda	boda	/dufo/	/borθion/
duende	bomba	/dumpo/	/braio/
duna	botín	/dunsla/	/briundo/
falda	brazo	/fafi/	/bromba/
faro	brillo	/farfa/	/buxor/
fase	burro	/fefad/	/burto /
feria	camión	/fende/	/kafin/
fiesta	cartón	/fersor/	/kareks/
final	centro	/fesdion/	/kaudel/
flauta	cepo	/fidio/	/θefa/
flora	cerdo	/fime/	/θembe/
folio	cerro	/finra/	/θemo/
frase	cesión	/firo/	/θenke/
fruta	césped	/fiuo/	/θepro /
funda	cetro	/fueda/	/θeurme/
fusil	ciclón	/fuinle/	/θibra/
ladrón	ciencia	/lade/	/θixon/
landa	cierre	/lanfio/	/θilo/
laurel	cifra	/larlon/	/θimpion/
lema	cine	/laron/	/θinre/
leña	ciprés	/lasil/	/θira/
líder	cita	/lelpo/	/θirfol/
lima	clase	/lirlen/	/θibad/
lino	clero	/lonse/	/koel/
lira	color	/luensa/	/koio/
litro	conde	/lulta/	/kona/
lomo	crema	/mefa/	/koθo/
luna	crimen	/melsa/	/kufa/
luto	cuadro	/menlo/	/kula/
madre	cuenca	/mepe/	/kunse/
maña	culpa	/meslo/	/kupia/
menú	cuota	/meua/	/kuspro/
metro	genio	/midra/	/xesur/
misa	germen	/mienlo/	/xikal/
misil	gestor	/mifa/	/xiespa/
mito	jaguar	/milda/	/xirka/

Appendix I – Stimuli lists of Experiment 3 (MEG)

molde	jarrón	/minsa/	/xanti/
momia	jefe	/misda/	/xarθa/
muerte	jota	/muenlo/	/xiasio/
multa	juego	/nanda/	/xiekel/
musa	jueza	/narla/	/xinlel/
naipe	jugo	/nefa/	/xola/
nieto	junco	/nofo/	/xoθis/
niño	jungla	/nopa/	/xuero/
nodo	junta	/nuno/	/xuxa/
norte	kaki	/nuro/	/xusa/
nutria	kayak	/paida/	/kade/
panel	kebab	/papor/	/kaifan/
parto	kiwi	/pefa/	/kalxa/
pastel	quema	/peia/	/kaua/
patio	queso	/penla/	/keɲa/
piano	quiebra	/perto/	/kifo/
pintor	quince	/pifa/	/kixil/
pista	quintal	/pifion/	/kibo/
plasma	vaca	/pinor/	/kire/
portal	vaina	/pirmud/	/bafria /
postre	vapor	/piula/	/baxed /
premio	vatio	/puelmo/	/bamio /
puente	vejez	/pulfo/	/basia/
puesta	verdad	/tamio /	/baslor/
tapia	verso	/tarna /	/berpa/
temor	viaje	/tefa/	/bienra/
templo	vidrio	/tere/	/bieθad/
tenis	vigor	/tifi/	/bila/
timón	violín	/tofa/	/bisma/
titán	virus	/tolo/	/borsal/
trapo	vista	/torfe/	/buerko/
trato	zarpa	/trela/	/bulθa/
trecho	zarza	/trira/	/θikro/
trompa	zueco	/troɲe/	/θiro/
trueno	zumba	/tuma/	/θorxe/
túnel	zurdo	/tusfo/	/θubo/

Appendix J – Stimuli lists of Reading Aloud Task

French – Words		**Spanish – Words**		**French – Pseudowords**		**Spanish – Pseudowords**	
Consistent	*Inconsistent*	*Consistent*	*Inconsistent*	*Consistent*	*Inconsistent*	*Consistent*	*Inconsistent*
adulte	filet	dominó	boda	avulme	ginais	posinó	cona
acheteur	fable	dios	broma	avedeur	jable	piom	crola
armure	festin	desierto	vampiro	artule	jaisetin	pesierno	kamniro
loutre	fruit	disparo	vacuna	moubre	stuie	pismaro	cajuna
bordure	phénomène	día	vestido	roldure	gébomène	piu	quesfido
bal	coffre	pera	calvo	dave	sogre	dema	valjo
douane	cave	pantera	cabra	bouale	sane	danlera	bazra
lame	carte	palma	convento	nabe	sarche	dalfa	bomtento
nounou	kiosque	pared	química	pouloux	ciofque	damed	vimiba
mur	quart	peto	kárate	tude	jamme	defa	bárale
match	cigarette	fuente	cereza	danne	quifarette	nuense	jereca
montagne	citrouille	fresa	cierre	tonmagne	quitrouze	luesa	rieje
marche	salade	flauta	zarcillo	ralche	catade	muauta	rarjillo
radar	savon	frontera	cita	badane	camon	tronlera	jilla
raton	salon	lente	joroba	bamon	cadon	fense	zoroza
chou	jongleur	loma	jardín	cheux	phoncleur	foña	rarfín
tour	judas	moto	joya	roube	fuva	lollo	zocha
tache	gel	molde	roca	mave	faine	folpe	joma
tarte	journal	medusa	rezo	varche	fourballe	feduta	ceco
neveu	geste	nadador	raíz	beteux	phecte	ladafor	zaíj

Appendix K– Stimuli lists of Picture Naming Task (referents)

French		**Spanish**	
Consistent	*Inconsistent*	*Consistent*	*Inconsistent*
bavoir	ferme	dado	barba
banane	fil	dama	bicho
barbe	fusée	dardo	boca
beurre	phoque	dedo	bosque
bourdon	photo	dálmata	burro
bouton	casque	delantal	vaca
bûche	corne	ducha	valla
dame	coq	duelo	vaso
double	couche	duende	vela
douche	crabe	duna	violín
doudou	kiwi	pala	cable
dortoir	koala	pan	caña
dune	queue	pato	coche
palme	quiche	pelo	codo
parachute	quille	percha	cuna
poumon	ceinture	parche	kilo
plume	cercle	pollo	kiwi
poire	cerise	pelota	koala
poche	ciseaux	pulpo	queso
pomme	citron	puño	quiosco
lama	sac	flecha	regalo
lune	scie	falda	regla
moule	seringue	farola	reloj
mouton	serpent	león	cebolla
voiture	singe	luna	cebra
robe	genou	nudo	ciervo
route	gilet	noche	zapato
table	jambe	melón	zorro
talon	jambon	montaña	jabón
vache	juge	maleta	jaula