eman ta zabal zazu

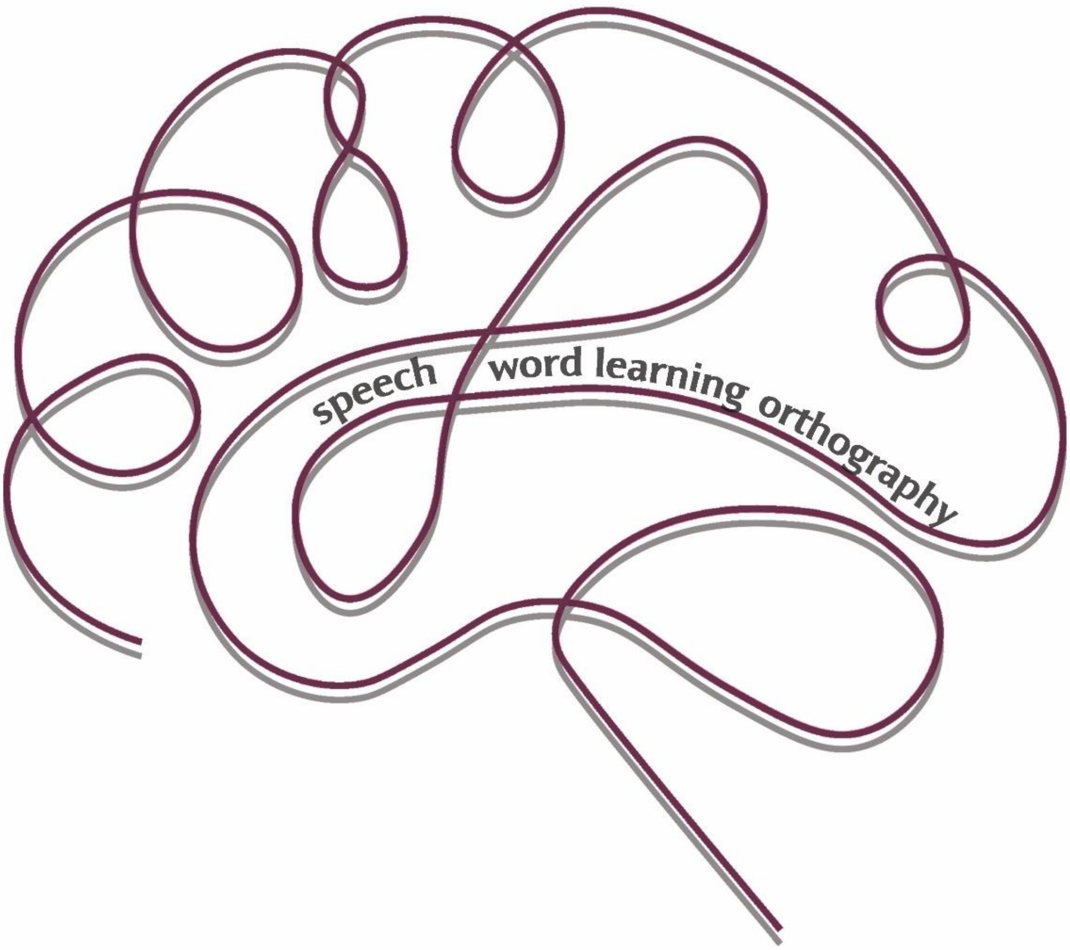
Universidad del País Vasco
Euskal Herriko Unibertsitatea

The role of orthography in auditory word learning

Doctoral thesis by Mina Jevtović

Supervised by

Prof. Clara D. Martin and Dr. Efthymia Kapnoula



Donostia – San Sebastián, 2022.

The role of orthography in auditory word learning

Doctoral thesis by

Mina Jevtović

Supervised by

Prof. Clara D. Martin and **Dr. Efthymia Kapnoula**

eman ta zabal zazu

Universidad del País Vasco
Euskal Herriko Unibertsitatea

NAZIOARTEKO BIKAINTASUN CAMPUSA
CAMPUS DE EXCELENCIA INTERNACIONAL

Donostia-San Sebastián, 2022

Basque Center on Cognition, Brain and Language
Paseo Mikeletegi 69
Donostia-San Sebastián, Spain
November 2022

The role of orthography in auditory word learning

Doctoral thesis by

Mina Jevtović

Supervised by

Prof. Clara D. Martin and **Dr. Efthymia Kapnoula**

Universidad del País Vasco Euskal Herriko Unibertsitatea

NAZIOARTEKO BIKAINTASUN CAMPUSA
CAMPUS DE EXCELENCIA INTERNACIONAL

BASQUE CENTER ON COGNITION, BRAIN AND LANGUAGE

Donostia-San Sebastián, 2022

This research was supported by a Predoctoral fellowship (associated to the project PSI2017 82941-P; grant number PRE-2018-083946) from the Spanish Ministry of Science, Innovation and Universities and the Fondo Social Europeo.

A six month research stay supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 317633480 – SFB 1287 was carried out at the University of Potsdam, Germany.

Open Science Statement

This thesis was written using the open-source R package Bookdown, and its online version is available at: https://minajevtovic.github.io/PhD_thesis/

All data, stimuli material, experimental and statistical scripts generated in the course of the present thesis are available at: https://osf.io/bhj9y/

За моју маму,
моју Наду.

Acknowledgements

It was the best of times,
it was the worst of times...
Charles Dickens

Much like Dickens' *A Tale of Two Cities* my PhD journey was full of ups and downs. This space is dedicated to people who celebrated my "ups" and helped me overcome the "downs".

Tout d'abord je tiens à remercier la personne qui a - bien qu'il m'ait fallu du temps à m'en rendre compte - vraiment toujours cru en moi. Clara, merci infiniment de m'avoir supportée ces trois dernières années (supportée dans tous les sens du terme; que ce soit en français ou en anglais). Merci pour ton temps et ta patience envers moi; pour toutes les fois où je voulais arrêter la thèse et où tu as réussi à me remonter le moral et me rappeler à quel point j'aimais ce travail. Même si le parcours n'a pas été très simple (ni pour toi ni pour moi), on s'est vraiment bien amusées toutes les deux ! Et en dépit de tous les obstacles auxquels on a dû faire face, on a réussi à réaliser pas mal de choses assez incroyables ensemble (comme par exemple achever cette thèse en avance et être les premières au BCBL à obtenir le POP !).

Dank, liebe Antje, für alles. Danke, dass du mich an meinem ersten Tag der Promotion herzlich empfangen hast und dass du dich immer um mich gekümmert hast und dir Zeit genommen hast. Auch habe ich mich immer über die Pralinen gefreut, die du mir aus Deutschland mitgebracht hast.

Huge thank you to all the new and former members of the Speech and Bilingualism group; thank you for all the feedback and support during the past three years. Many thanks as well to Arty Samuel for all the research and career-related advice! The biggest thank you (in Greek!) to the best second supervisors Effie and Alexia! You both always knew how to make me smile even on my worst days. I will forever be grateful to both of you for looking out for me during my PhD years.

Thank you to the external reviewers Dr. Manon Jones and Prof. Dr. Niels Schiller, as well as the committee members Prof. Dr. María Teresa Bajo Molina, Prof. Dr. François-Xavier Alario and Dr. Brendan Costello for the time taken to evaluate the work I have done for this thesis!

Huge Eskerrik asko a toda la gente del lab! Amets, Itziar no lo podría haber hecho sin vosotras. Gracias por cuidarme desde mis primeros días en el BCBL. Ainhoa, Larraitz, Mamen - trabajar con cada una de vosotras ha sido un enorme placer. ¡Vaya donde vaya siempre os voy a echar de menos! Gracias también a todas las *admin girls*, Ana, Eider, Larraitz y Maider, por la paciencia y la inmensa ayuda con todo el papeleo y la burocracia; sobre todo durante los últimos meses de mi doctorado!

Many thanks to several people from Donosti who made one of the loneliest periods of my life feel a bit less lonely. Hana, my balkan buddy, puno ti hvala što si uvek bila pravi prijatelj! Hvala i divnoj porodici Kodro na nekoliko predivnih susreta koji su me podsetili šta znači toplina doma. Romain, merci d'avoir été mon ami lors des moments les plus difficiles de ma vie ! Merci aussi à tous les autres membres du groupe French lunch (Svetlana, Tiphaine et Florent); merci beaucoup de m'avoir acceptée parmi vous - bien que j'aie jamais vraiment mangé à la même heure que vous...
Gracias también a todos los predocs del "piso" por las risas y los buenos momentos.

A special *danke* to all the wonderful people from the University of Potsdam for giving me the best 6 months of my entire PhD experience! Thank you to the SFB1287 for funding my research stay and for giving me the opportunity to work, grow and learn at such a wonderful and inspiring place (with the best Mensa in the world!). Thank you to my dear office mates Melina and Zirui for making our office such a happy place. To Raúl, Isa and Michaela for showing me that moving to a new country doesn't have to be scary and stressful. Finally, thank you to all the people from the CSLM lab who made me feel welcome from day one! Marie, Shereen, Solveig and Paula, thank you for being my friends and accepting me as I am! Pam thank you for being the best *academic sister* anyone can dream to have! Thank you for all the support, advice, jokes and PhD memes! You and Garrett brought so much light and hope into my life at the moment when everything seemed grey! Audrey, merci de m'avoir redonné le goût à la vie académique ! Si j'ai pu terminer cette thèse avec succès c'est en énorme partie grâce à toi ! Danke für alles!

Finally, my deepest gratitude goes to my family and friends who were my biggest support from day one. Ksenija hvala ti što me podržavaš i bodriš od treće godine osnovnih studija! Teško je (lako nije!) bilo dogurati do ovde. Andrea, thank you for being my person već više od 20 godina! Je t'aime do neba i nazad. Sara gracias infinitas por estar a mi lado desde el primer día del máster! Por apoyarme y por no dejar que me rinda cuando lo quería dejar todo. Pero sobretodo, gracias por ser mi familia en España! Ciara, thank you for always being there for me and for cheering me up even from afar! Hvala puno mojoj dragoj sestri Ani Ivanović na najlepše dizajniranoj naslovnoj strani.

Mama, tata, Ćiki sad sam i zvanično najpametnija u porodici (posle Tea naravno)! Volim vas najviše na svetu. Hvala vam za sve!

Abstract

Learning to read affects spoken language processing. It has been shown for instance, that the way words are written can facilitate or hinder their auditory perception: words containing rhymes with unique spellings (e.g., /oʊb/ in globe can only be written as <obe> in English) are recognised faster than those containing rhymes with multiple possible orthographic representations (e.g., /eɪm/ in name can be written as either <ame> or <aim>; see Ziegler & Ferrand, 1998). The present thesis expands on the relationship between spoken and written language by investigating whether orthography plays a role in spoken word learning. Recent findings have shown that both children and adults use their knowledge of sound-to-letter correspondences to generate preliminary orthographic representations (hereafter *orthographic skeletons*) for aurally acquired words (Wegener et al., 2018; Wegener, Wang, Nation, & Castles, 2020). Consequently, we set out to further investigate the mechanism(s) by which sounds are mapped onto letters during spoken word learning.

In Experiment 1, we tested whether orthographic skeletons are generated for all aurally acquired words (regardless of the number of possible spellings) or only for words with unique, and hence completely predictable spellings. Forty-eight Spanish adult speakers first acquired a set of novel Spanish words through aural training, and then saw their spellings in a self-paced reading task. Crucially, words were shown in their unique (consistent words) or one of their two possible spellings (preferred and unpreferred inconsistent words). Participants' spelling preferences for all novel words were collected two weeks before the experiment took place. The results showed no differences in reading times between preferred and unique spellings. However, words with unpreferred spellings yielded significantly longer reading times as compared to words with unique spellings. This indicates that participants indeed generated orthographic skeletons for novel spoken words even when there was uncertainty regarding the correct spelling. In Experiment 2 we further investigated the leniency of the process underlying the creation of orthographic skeletons by testing speakers of an opaque language. Compared to Spanish, in which most sounds map onto only one grapheme thus making Spanish highly consistent for spelling, French is highly inconsistent for spelling. In French one sound is usually associated with more than one letter (e.g., the sound /o/ has at least three orthographic representations <o>, <au> and <eau>). Due to the overall higher probability of generating incorrect orthographic representations, French speakers may not engage in the process of generating orthographic skeletons when multiple spelling options are possible. Nevertheless, the pattern of results was similar to the one observed with Spanish speakers. This suggests that generating orthographic skeletons during spoken word learning may be beneficial independently of language opaqueness. Finally, to test whether orthographic skeletons are generated every time a new spoken word is acquired (i.e., automatically) or only when generating them is beneficial for the task (i.e., strategically), two groups of Spanish speakers completed the same aural training task, but received different instructions. While *active learners* knew they were supposed to learn the words, *passive learners* were naive regarding the aim of the experimental task. As in Experiment 1, we observed longer reading times for previously acquired spoken words with unpreferred spellings, but only in the group of active learners. This suggests that generating orthographic skeletons during spoken word learning may not be automatic, but can actually represent a strategy participants employ with the aim to facilitate the learning process.

Overall, the findings of the present thesis suggest that skilled readers can use their knowledge of sound-to-letter mappings to generate preliminary orthographic representations for aurally acquired words before encountering them in writing. Importantly, we show that orthographic skeletons are generated even when there is uncertainty regarding the correct spelling, both in languages with overall high, as well as those with overall low probability of generating an incorrect representation. These findings raise a possibility that generating orthographic skeletons may be beneficial.

Finally, our data suggest that the process by which orthographic skeletons are generated might not be automatic and inherent to the cognitive system, but can actually be driven by strategies participants employ in order to facilitate the word learning process. Taken altogether, the present findings reveal new ways in which orthography affects spoken language processing.

Resumen en Castellano

La lectura y la escritura constituyen las principales herramientas de la humanidad para recopilar, almacenar y comunicar conocimientos a través del espacio y el tiempo. Además de su importancia en la evolución y el desarrollo de la especie humana, las habilidades de lectura y escritura son cruciales para el desarrollo del individuo. Aprender a leer permite acceder a la educación y, en consecuencia, a mejores oportunidades profesionales y sociales. La investigación en psicolingüística, y en la ciencia cognitiva en general, se ha interesado especialmente en cómo puede repercutir el aprendizaje de la lectura y la escritura en otras funciones cognitivas. Por ejemplo, se ha dedicado una gran cantidad de estudios a investigar los efectos que tiene la adquisición de la alfabetización en el procesamiento del lenguaje oral. Teniendo en cuenta que se ha demostrado que una buena capacidad de lectura está relacionada con un mejor procesamiento fonológico (véase Snowling y Hulme, 2021), los investigadores se han preguntado si el aprendizaje de la lectura puede influir de alguna manera en la percepción del lenguaje oral. Para ello, han aprovechado el hecho de que, mientras que algunos sonidos tienen representaciones ortográficas únicas (por ejemplo, el sonido /p/ solo puede escribirse con la letra <p>), los demás pueden corresponderse con más de un grafema (por ejemplo, el sonido /k/ tiene tres representaciones ortográficas <c>, <qu> o <k>). Muchos estudios emplearon esta (in)consistencia ortográfica para comprobar los efectos de la alfabetización en la percepción del habla. Lo que demostraron es que las palabras orales con rimas que pueden escribirse de una sola manera (por ejemplo, /oʊb/ en la palabra inglesa <globe>) se reconocen más rápidamente que las que contienen rimas con múltiples representaciones ortográficas posibles (por ejemplo, /eɪm/ en inglés puede escribirse como <eim> o <aim>). Cabe destacar que se han demostrado los mismos efectos en varios idiomas: inglés (Ziegler et al., 2008), francés (Ziegler & Ferrand, 1998), portugués (Ventura et al., 2004); así como en lectores menos competentes como los niños que están en las primeras etapas de la

adquisición de la alfabetización (Ventura et al., 2004). Estos resultados mostraron, por tanto, que la ortografía de las palabras conocidas afecta a su reconocimiento oral. Una de las explicaciones de los hallazgos observados que se proponen en la literatura es que, al escuchar una palabra nueva, su representación ortográfica se activa automáticamente, lo que afecta a su tiempo de procesamiento (Chéreau, Gaskell y Dumay, 2007). La literatura precedente muestra, por tanto, que las representaciones ortográficas ya existentes afectan al procesamiento del habla, lo que proporciona una fuerte evidencia de que la adquisición de la alfabetización influye en el procesamiento del lenguaje hablado. La presente tesis va un paso más allá al comprobar si las representaciones ortográficas pueden generarse incluso en ausencia de ortografía. En concreto, intentamos comprobar si la ortografía desempeña un papel en el aprendizaje de palabras orales.

La presente tesis

El objetivo de la presente tesis es seguir explorando las formas en las que el conocimiento ortográfico afecta al procesamiento del lenguaje hablado. Más concretamente, nos propusimos comprobar si la ortografía desempeña un papel en el aprendizaje de palabras orales nuevas. Pruebas recientes han mostrado que los niños de habla inglesa (Wegener et al., 2018; 2020), así como los adultos angloparlantes (Beyersmann et al., 2021), pueden utilizar su conocimiento de las correspondencias entre sonidos y letras para formar representaciones ortográficas preliminares (en adelante, *esqueletos ortográficos*) para las palabras que adquieren auditivamente. La presente tesis se basa en esta explicación, conocida como la *Hipótesis del Esqueleto Ortográfico*, e investiga el mecanismo responsable de las conversiones de sonido a letra durante el aprendizaje de palabras orales.

En su estudio original, Wegener y sus colegas (2017) enseñaron a niños de 9-10 años de habla inglesa el significado y la pronunciación de unas palabras nuevas. A continuación, midieron el movimiento ocular de los niños mientras leían oraciones que contenían grafías de palabras que habían adquirido previamente junto con las grafías de palabras que no les habían enseñado. Es importante destacar que, aunque la mitad de las palabras aparecían con su ortografía predecible (por ejemplo, /neʃ/ escrita como <nesh>), la otra mitad tenía una ortografía impredecible (por ejemplo, /kɔɪb/ escrita como <koyb> en lugar de <coib>). El patrón de resultados que observaron indicaba que los niños generaron esqueletos ortográficos para las palabras nuevas antes de verlas impresas. Concretamente, los autores observaron una lectura más fácil para las palabras que habían practicado auditivamente en comparación con las que no habían practicado, y eso se dio en las cuatro medidas de seguimiento ocular: produjeron tiempos totales de lectura más cortos, tiempos más cortos con la mirada fija, tiempos más cortos en la primera fijación, así como menos regresiones en las palabras con ortografía predecible. Además, se observó una interacción significativa entre el entrenamiento y la ortografía, lo que demuestra que solo se produjo una lectura más fácil en el caso de las palabras que habían adquirido previamente y que aparecieron con su grafía más predecible. Esos resultados se interpretaron de la siguiente manera: dado que las expectativas que los niños habían generado para las palabras predecibles que habían practicado coincidían con sus grafías reales, su lectura posterior fue más fácil. Por el contrario, como había una discordancia entre los esqueletos ortográficos que los niños habían generado para las palabras impredecibles que habían practicado y sus grafías reales, no tuvieron una mayor facilidad. Aunque proporcionan una sólida evidencia para la hipótesis del esqueleto ortográfico, estos resultados no dejan claro si los esqueletos ortográficos se generan para todas las palabras adquiridas a través de instrucción auditiva o solo para las palabras con ortografías altamente predecibles. Es decir, no queda claro si el hecho de no tener una mayor facilidad con las palabras previamente adquiridas con ortografías impredecibles se debe a que los niños generan esqueletos ortográficos incorrectos o a que no generan ninguna representación. Dado que estas palabras tenían ortografías impredecibles y, por tanto, múltiples, podría darse el caso de que los niños ni siquiera iniciaran el proceso de generar expectativas ortográficas cuando la incertidumbre respecto a las posibles grafías era alta. Además, los propios autores señalan que, debido a la complejidad del sistema de escritura del inglés, las ortografías predecibles e impredecibles no coincidían, entre otras cosas, con las frecuencias de bigramas. Por lo tanto, las propiedades del material utilizado podrían haber influido en el patrón de resultados que obtuvieron.

Así, en el experimento 1 nos propusimos investigar la permeabilidad del mecanismo que se encarga de generar esqueletos ortográficos. En concreto, comprobamos si los esqueletos ortográficos se generan incluso cuando, debido a las múltiples opciones ortográficas posibles, hay incertidumbre respecto a la ortografía correcta. Con ese fin, 48 adultos castellanoparlantes aprendieron, primero, unas palabras nuevas mediante instrucción auditiva y, luego, vieron cómo se escribían en una tarea en la que leían a su propio ritmo. Es importante destacar que las palabras nuevas se presentaban con su única ortografía (consistente) o con una de sus dos posibles ortografías (palabras inconsistentes preferidas y no preferidas). Las preferencias ortográficas de los participantes se recogieron dos semanas antes del entrenamiento auditivo. Solo se observaron tiempos de lectura significativamente más largos para las palabras previamente adquiridas con grafías no preferidas, mientras que la lectura de palabras con grafías preferidas no difería de la lectura de palabras con grafías únicas. Entendimos estos resultados como prueba de que los participantes habían, efectivamente, generado representaciones ortográficas para todas las palabras: aquellas que tenían una ortografía única y aquellas que tenían dos posibles. Curiosamente, un análisis exploratorio no planificado reveló una significativa correlación positiva entre la tendencia individual a generar esqueletos ortográficos (calculada como la diferencia en la lectura de palabras con grafías no preferidas y con grafías únicas) y el posterior recuerdo de las palabras que se muestran con sus grafías no preferidas. Esta correlación indicaba que los participantes más propensos a generar esqueletos ortográficos también recordaban mejor las palabras que se mostraban con su grafía impredecible. Los datos del experimento 1 se suman, así, a la hipótesis del esqueleto ortográfico al mostrar que, incluso cuando hay más de una ortografía posible, se generan representaciones ortográficas preliminares de palabras orales recién adquiridas. Sin embargo, estos resultados puede que estén limitados a los hablantes de lenguas relativamente consistentes como el castellano. De hecho, los hablantes de castellano rara vez se enfrentan a grafías que no coinciden con sus expectativas. Por lo tanto, podrían ser más proclives a generar esqueletos ortográficos incluso cuando hay riesgo de error.

Al contrario que el español, el sistema de escritura francés es muy inconsistente en cuanto a la ortografía. Dado que muchos sonidos tienen más de una representación ortográfica, predecir la ortografía correcta de una palabra oral nueva conlleva un mayor grado de incertidumbre y, por tanto, un mayor riesgo de error. Podría darse el caso de que los hablantes de lenguas inconsistentes, como el francés, no generen esqueletos ortográficos para palabras con más de una grafía altamente predecible. Por lo tanto, en el experimento 2 investigamos si se generan esqueletos ortográficos para palabras con más de una grafía posible incluso en una lengua en la que la probabilidad de generar una representación incorrecta es mayor. Para eso, 46 adultos francófonos completaron el mismo experimento en el que, primero, aprendían palabras nuevas orales en francés con una o dos grafías posibles, y luego las veían por escrito en una tarea de lectura, a su ritmo. Además, nos preguntamos si el vínculo que habíamos observado previamente entre la generación de esqueletos ortográficos puede estar modulado por otro constructo que ha demostrado ser importante para el aprendizaje de palabras orales, es decir, la memoria fonológica a corto plazo (PSTM, phonological short-term memory; véase Gathercole, 2006; Gathercole y Baddeley, 1989). Teniendo en cuenta que se ha demostrado que el vínculo entre ambos disminuye con la edad (Gathercole, 2006), quisimos explorar si podría deberse a la aparición del mecanismo de conversión de sonidos a letras durante el aprendizaje de palabras orales. Medimos la capacidad de PSTM de los participantes mediante una tarea de repetición de pseudopalabras. Los resultados del experimento 2 mostraron que los esqueletos ortográficos de las palabras con ortografía incierta se generan incluso en idiomas en los que la probabilidad de error es alta. En primer lugar, los participantes franceses leyeron más rápido las palabras que habían practicado previamente que las que no habían practicado. Es importante destacar que, mientras que no observamos diferencias en los tiempos de lectura entre las palabras entrenadas únicas y las preferidas, las grafías no preferidas de las palabras inconsistentes practicadas dieron lugar a tiempos de reacción significativamente más largos. Curiosamente, demostramos que el vínculo previamente observado entre la generación de esqueletos ortográficos y el recuerdo de palabras está efectivamente modulado por la capacidad de PSTM. La interacción entre la tendencia individual a generar esqueletos ortográficos y las puntuaciones de PSTM en la exactitud a la hora de recordar palabras reveló que solo los participantes con puntuaciones de PSTM de medias a altas pudieron beneficiarse de la generación de esqueletos ortográficos durante el aprendizaje de palabras orales. Este hallazgo podría explicar el descenso observado en el vínculo positivo entre el PSTM y el aprendizaje de palabras orales que se produce con la edad (Gathercole, 2006). Una vez que se han adquirido las reglas de conversión de sonido a grafía, tanto los niños como los adultos podrían confiar en ellas para facilitar la adquisición de nuevas palabras orales. En definitiva, los resultados del experimento 2 sugieren que la generación de esqueletos ortográficos puede ser beneficiosa para el sistema cognitivo.

Los resultados de los dos primeros experimentos muestran, por tanto, que los adultos que son lectores hábiles generan esqueletos ortográficos para las palabras orales recién adquiridas, incluso cuando hay incertidumbre respecto a la grafía correcta. No obstante, en estos dos experimentos, así como en los realizados por Wegener y sus compañeros (2018, 2020), los participantes adquirieron palabras nuevas mediante una instrucción explícita. Es decir, sabían que después se les iba a evaluar. Como se ha demostrado que la ortografía tiene un efecto positivo en el aprendizaje de palabras nuevas y que puede utilizarse, de hecho, como herramienta mnemotécnica durante el proceso de aprendizaje (Johnston et al., 2004; Rosenthal y Ehri, 2008), podría darse el caso de que generar esqueletos ortográficos fuera el resultado de un proceso estratégico que los participantes aprovecharan conscientemente con el objetivo de facilitar el proceso de aprendizaje. En el experimento 3 investigamos si los esqueletos ortográficos se generan inconscientemente, como una respuesta automática al input fonológico nuevo durante el aprendizaje de palabras orales nuevas, o si la generación de esqueletos ortográficos representa una estrategia que ayuda a los participantes a aprender palabras orales nuevas con más facilidad. Con ese objetivo, evaluamos a dos grupos de participantes que tuvieron que aprender palabras orales nuevas de forma activa (explícita) o pasiva (implícita). Más concretamente, ambos grupos de participantes realizaron la misma tarea de entrenamiento auditivo, pero mientras que a los que aprendían activamente se les indicaba explícitamente que retuvieran los nombres de los objetos nuevos, los que lo hacían de manera pasiva no sabían cuál era el objetivo de la tarea. Igual que en los dos primeros experimentos, a los participantes se les enseñó la ortografía de las palabras en una tarea en la que tenían que leer unas frases a su ritmo. En general, los resultados del experimento 3 indican que mientras los lectores activos utilizaron su conocimiento de las correspondencias sonido-letra para generar esqueletos ortográficos, los que aprendían pasivamente no lo hicieron. Como los tiempos de lectura observados en el experimento 1 se reprodujeron solo en el grupo de lectores activos, y no en el grupo de aprendizaje pasivo, consideramos que es una evidencia de que generar esqueletos ortográficos es el resultado de un proceso estratégico y no de un proceso automático de conversión de sonido a letra. Además de proporcionar mayor entendimiento sobre el proceso que conduce a la creación de esqueletos ortográficos, los datos del experimento 3 se suman a la literatura precedente al mostrar que los efectos ortográficos en el procesamiento del habla no siempre son automáticos.

Debate general

Los resultados de la presente tesis tienen implicaciones tanto para las teorías de la lectura como para las del aprendizaje de palabras nuevas. En primer lugar, mostramos que se pueden generar representaciones ortográficas en ausencia de ortografía, incluso cuando hay riesgo de generar una representación incorrecta. Este hallazgo pone de manifiesto la importancia y la prevalencia de la lectura en nuestras vidas, así como nuestra tendencia a vincular la ortografía al lenguaje oral. Debemos destacar que la tesis se suma a la literatura precedente al revelar que las consecuencias de generar representaciones ortográficas preliminares para la posterior lectura dependen de las propiedades de los sistemas de escritura. Crear esqueletos ortográficos en lenguas con correspondencias entre sonidos y letras complejas, como el francés y el inglés, conduce a una mayor facilidad en la lectura al enfrentarse a grafías correctamente predichas. Por el contrario, la generación de esqueletos ortográficos en lenguas con correspondencias de sonido a letra simples conlleva una desventaja en la lectura de grafías predichas de manera incorrecta. Esto sugiere que generar esqueletos ortográficos puede ser más beneficioso en lenguas opacas como el inglés y el francés. En consonancia con la conclusión de que la generación de esqueletos ortográficos durante el aprendizaje de palabras orales es beneficiosa, también tenemos los resultados de los dos análisis exploratorios que estudian el recuerdo de palabras al final del experimento. En concreto, tanto los participantes españoles como los franceses con mayor tendencia a generar esqueletos ortográficos recordaban mejor las palabras orales recién adquiridas. Este hallazgo también revela que existen importantes diferencias individuales en la generación de esqueletos ortográficos. Además, mostramos que esta influencia de la ortografía en el aprendizaje de palabras orales puede estar impulsada por estrategias personales de los participantes. Se ha demostrado que la ortografía es de ayuda durante la adquisición de vocabulario nuevo (Ehri y Wilce, 1979). Añadimos a ese conocimiento que los lectores expertos pueden hacer un uso voluntario de su conocimiento ortográfico (el conocimiento de las correspondencias de sonido a ortografía) para generar pistas adicionales para la memoria en forma de esqueletos ortográficos durante el proceso de aprendizaje. Por último, el presente trabajo tiene también implicaciones prácticas para la lectura y el aprendizaje de palabras nuevas. En primer lugar, dado que no tienen nada que perder, tanto a los adultos como a los niños que aprenden a leer en una lengua inconsistente se les debería animar a generar representaciones ortográficas para todas las palabras auditivamente familiares. Esta sugerencia podría incorporarse a las intervenciones de lectura ya existentes, así como a los programas de aprendizaje de idiomas. Además, a pesar de que todos los estudios que investigan el proceso de generación de esqueletos ortográficos se han realizado con hablantes que leían en su primera y más dominante lengua (L1), sus resultados también tienen implicaciones en la adquisición de una segunda lengua. La adquisición de habilidades lingüísticas orales en la lengua materna casi siempre es anterior a la adquisición de la lectura y la escritura. En cambio, en la(s) segunda(s) lengua(s) (L2), que se adquiere(n) más tarde en la vida, y la mayoría de las veces en un contexto de clase, las dos habilidades suelen desarrollarse en paralelo. Por lo tanto, podría darse el caso de que los efectos ortográficos en el aprendizaje de palabras orales sean más fuertes en una L2, ya que las habilidades orales y escritas alcanzan un nivel más equilibrado en las lenguas adquiridas en un aula. Por lo tanto, en futuros estudios se podría investigar el proceso de generación de esqueletos ortográficos en las lenguas adquiridas más tarde.

Conclusión

En general, los resultados de la presente tesis sugieren que los lectores expertos pueden utilizar su conocimiento de las correspondencias entre sonidos y letras para generar representaciones ortográficas preliminares de las palabras adquiridas auditivamente antes de verlas por escrito. Y, lo que es más importante, lo hacen incluso cuando, debido a las múltiples opciones ortográficas posibles, no hay certezas respecto a cuál será la grafía correcta. El trabajo presentado aquí muestra, además, que los esqueletos ortográficos de las palabras con múltiples grafías se generan tanto en las lenguas con una probabilidad alta de generar una representación incorrecta (por ejemplo, el francés), como en aquellas con una probabilidad baja (por ejemplo, el español). Por último, mostramos que el proceso por el que se generan los esqueletos ortográficos puede que no sea automático e inherente al sistema cognitivo, sino que podría estar impulsado por estrategias que los participantes emplean para facilitar el proceso de aprendizaje de palabras. En conjunto, los presentes hallazgos se suman a la literatura precedente al revelar nuevas formas en las que la adquisición de la lectura afecta al procesamiento del lenguaje hablado.

Contents

Introduction

Apprendre à lire, c'est allumer du feu;
toute syllabe épelée étincelle.
Victor Hugo

Reading and writing represent humankind's principal tools for transmitting and retrieving information through space and time. The fact that you are reading this thesis, which was written at a certain point in time with a very precise aim, demonstrates the power conveyed by only 26 visual symbols (i.e., letters) ordered in a specific way. However, compared to spoken language, the system that enabled the development of reading and writing (i.e., orthography) is a *relatively* new invention in human evolution. As such, its role has long been neglected in the scientific study of language. Namely, despite its undisputable importance in collecting, storing and communicating knowledge, it seems almost impossible that some of the most prominent linguists of the 20$^{\text{th}}$ century did not even consider orthography to be a part of language (Rutkowska, 2017). According to Bloomfield for instance, orthography as an object of research does not belong to linguistics, and should therefore be completely removed from its scientific inquiry. Considered as being one of the leading voices responsible for the disregard of orthography in the scientific study of language, Bloomfield viewed written language only as a tool which was to be used to represent speech:

> "Writing is not language, but merely a way of recording language by means of visible marks" (Bloomfield 1933: 21).

Similar views in line with this *relativistic* school of thought which postulates the complete reliance of written on spoken language, were shared by Saussure, who considered orthography a secondary system, entirely dependent on speech:

> *"Langue et écriture sont deux systèmes de signes distincts ; l'unique raison d'être du second est de représenter le premier"*

> [Language and writing are two distinct systems of signs; the second exists for the sole purpose of representing the first] (Saussure 1989: 45).

as well as Sapir, who viewed written language as a circulating medium used to express speech:

> "The written forms are secondary symbols of the spoken ones — symbols of symbols — yet so close is the correspondence that they may, not only in theory but in the actual practice of certain eye-readers and, possibly, in certain types of thinking, be entirely substituted for the spoken ones." (Sapir 1921: 20).

This view has later been abandoned in favour of a more *autonomous* status of orthography according to which writing does more than *just* represent speech (see Liuzza, 1996). Notwithstanding, the superiority of speech in relation to spelling[1] can also be seen at the initial stages of psycholinguistic research of written language processing (Rastle, 2019). This *primacy* of speech over reading and writing, and in particular the complete reliance of written on spoken language, is best reflected in the fact that (silent) reading was initially seen as speech happening inside one's head (i.e., inner speech; see, Huey, 1908). Moreover, even in the modern study of reading, the notion of "mental pronunciation" or phonological decoding, plays a central role in understanding how letters are converted into sounds (Rozin & Gleitman, 1977). With this in mind, and considering that speech (spoken language), precedes reading and writing (written language), not only phylogenetically (in the course of human evolution) but also ontogenetically (in the course of individual development), it is not surprising that mainly effects of spoken on written language have been studied in the psycholinguistic literature. We therefore start this

[1]Note that the term *orthography* denotes a more broader term than *spelling*. Namely, while spelling comprises phoneme-to-grapheme conversion rules, the term orthography refers to both spelling as well as the capitalization, punctuation, and word division allowed in a certain language. Given that the distinction between the two is usually not (explicitly) made in the psycholinguistic research, for the sake of consistency, in the present thesis spelling and orthography will also be used as synonyms.

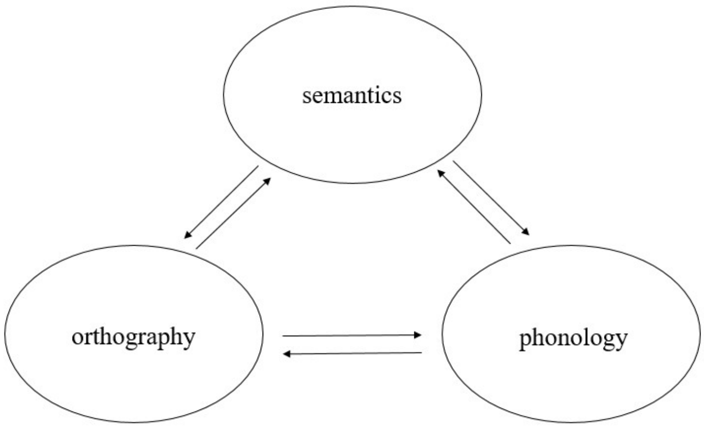
introduction by presenting some of the key findings showing how our experience with spoken language affects written language processing (i.e., effects of phonology on orthographic processing). Note however that the focus of the present thesis is actually on the reverse effects (i.e., effects of orthography on phonological processing).

The impact of phonology on written language processing

Examples of how spoken language (i.e., phonology) impacts written language processing (i.e., visual word recognition or silent reading) can be found in everyday reading contexts. For example, skilled readers tend to overlook orthographic errors that match the phonology of the target word (i.e., pseudowords that are phonetically identical to a real word such as crane/crain) and accept these pseudohomophones as real words during reading. Similarly, they may need more time to read and comprehend sentences containing sound repetitions (i.e., tongue-twisters) even when reading “in their heads” (i.e., during silent reading). This anecdotal evidence is nevertheless supported by a vast amount of experimental findings showing that phonology is activated during written language processing (Berent & Perfetti, 1995; Lukatela & Turvey, 2000; Pexman, Lupker, & Reggin, 2002; for a review, see Brysbaert, 2022). As in the anecdotal example described above, the studies in question made use of homophones - words that sound the same (i.e., share phonology) but differ in their meanings and spellings (e.g., <rows> and <rose>) - to probe for the activation of phonology during written language processing. For example, Van Orden (1987) showed that skilled readers need more time to detect homophone errors in a semantic decision task. Participants in his study more often incorrectly classified MEET (homophone of MEAT) as a type of food as compared to MELT. As this *homophony effect* was present regardless of the orthographic similarity between the correct and the homophone word, the author argued that the effects were indeed driven by automatic activation of phonological information during early visual word recognition.

Early and automatic activation of phonology in visual word recognition is further supported by findings showing faster recognition of words preceded by visual primes with the same pronunciation and similar spelling (e.g., <mayd> before <MADE>) as compared to primes with similar spellings only (e.g., <mard> before <MADE>; Perfetti, Bell, & Delaney, 1988). These homophone effects, replicated in a variety of experimental tasks and paradigms, such as letter detection (Drewnowski & Healy, 1982) or silent-proofreading task (Daneman & Stainton, 1991), were explained in the light of interactive models of reading and visual word recognition. According to these models, which all assume bidirectional links between orthographic and phonological codes (see Figure I.1), different representational levels interact during visual word recognition (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; e.g., Plaut, McClelland, Seidenberg, & Patterson, 1996). Due to this interactive flow of activation, phonological representations get automatically activated by their orthographic counterparts during written word processing, and that regardless of their importance for the task at hand.

Figure I.1: The Triangle Model of Word Reading



The triangle model of word reading adapted from Plaut, Seidenberg, McClelland, and Patterson (1996) assuming bi-directinal links between phonological, orthographic and semantic levels of representation.

Further evidence supporting the automatic activation of phonology in written language processing can be found in research done with bilinguals and second language learners. For instance, Friesen and Jared (2012) showed that bilinguals need more time to process written words that share the sound but not the spelling or the meaning across their two languages (i.e., interlingual homophones; e.g., the word /ʃuː/ written as 'shoe' which means cabbage in French) as compared to words specific to one of their languages. Cross-linguistic homophone effects were also found in Dutch-French bilinguals who showed similar priming effects when prime and target belonged to different languages (i.e., interlingual priming effects) as compared to when they belonged to only one (Brysbaert, Van Dyck, & Van de Poel, 1999). Importantly, these results hold regardless of the proficiency in a second language (Duyck, Diependaele, Drieghe, & Brysbaert, 2004).

Finally, research looking into readers with reading disorders such as developmental dyslexia, highlights the importance of phonology and phonological skills in written language processing (Saksida et al., 2016; for a review, see Snowling & Hulme, 2021). Given that readers with dyslexia show poor performance in tasks requiring phonological processing (e.g., phonological awareness tasks; Snowling, 1995), and that many reading researchers view issues with phonological processing as the core deficit of this reading disorder (for a review, see Snowling & Hulme, 2021), phonology is considered as being indispensable for successful processing of written language.

In summary, previous research shows pervasive effects of phonology on orthographic processing. It seems that our experience with spoken language cannot be *shut down* even when it is irrelevant for the task, as is the case during written language processing. Prominent models of reading (Coltheart et al., 2001) and visual word recognition (Plaut et al., 1996) all assume the existence of bidirectional links (as well as the flow of activation) between phonological and orthographic lexicon. In addition, written communication is nowadays just as important as speech. Thus, it was only a matter of time when the reverse effects, specifically, the effects of orthography on spoken language processing, would start being explored in the literature. The present thesis pertains to this line of research as it explores indirect effects of orthography on spoken word processing, and in particular, spoken word learning.

The impact of orthography on spoken language processing

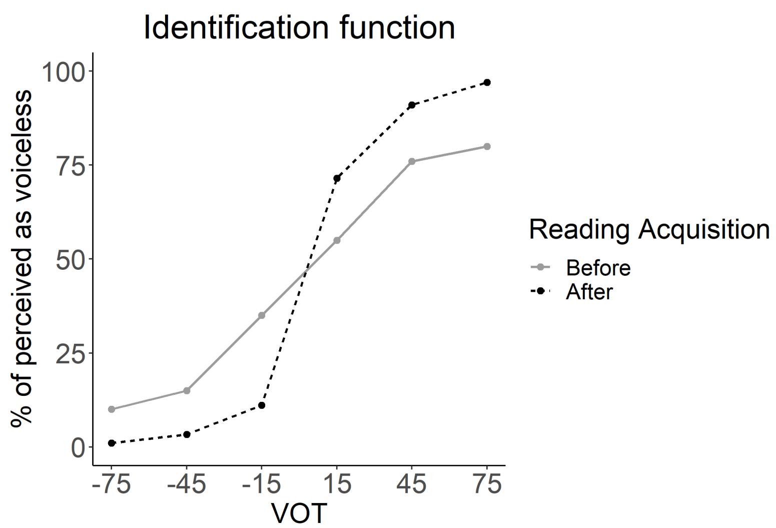
The idea that orthography is involved in spoken language, and in particular, phonological processing, is closely tied to the notion of phonological awareness and the development of metalinguistic skills. Morais, Alegria and Bertelson (1979) were one of the first to reveal a link between phonological awareness and reading ability. They compared the performance of age-matched literate and illiterate adults in tasks involving addition and deletion of sounds from spoken words and nonwords. They showed that literates completed both tasks easily (i.e., the percentage of accuracy was close to ceiling when real words were used, and relatively high for nonwords), whereas accuracy in the group of illiterates was rather low. Following up on these findings, numerous studies reported a positive link between reading skills and the performance on phonological awareness tasks (Saksida et al., 2016; for a review, see Snowling & Hulme, 2021). In consequence, phonological awareness, defined as a person's ability to perceive, think about, and manipulate individual speech sounds (phonemes) within spoken words (Snowling & Hulme, 1994), has been considered as being dependent on the acquisition of reading (Morais & Kolinsky, 2017; but see for example, Lundberg, Frost, & Peterson, 1991). Evidence showing that phonological awareness is indeed dependent on orthographic knowledge can be found in tasks in which relying on orthography actually comes with a processing disadvantage (Castles, Holmes, Neath, & Kinoshita, 2003; for similar results, see Tyler & Burnham, 2006). Castles and colleagues (Castles et al., 2003) manipulated the transparency of the to-be-removed sounds, and observed worse deleting performance for words containing sounds with opaque sound-to-letter correspondences (e.g., removing /n/

from <knuckle>) as compared to those with transparent links (e.g., removing /b/ from <buckle>). The debate on whether there is a causal relationship between phonological awareness and reading success is still ongoing in the literature. Nevertheless, there seems to be agreement in that the two skills develop in interaction and that the relationship between them is reciprocal (for a review, see Castles & Coltheart, 2004). In the reminder of this section we present evidence showing how learning to read affects speech processing. We will first review studies showing reading related differences in individual speech sound processing. Since the focus of this thesis is on spoken words, we will then move on to work reporting orthographic effects in spoken word processing.

The impact of orthography on speech sound processing

More evidence supporting the idea that reading acquisition enhances the perception of speech in terms of speech sounds is found in studies showing improvements in speech perception overlapping in time with the official onset of reading instruction. For instance, Hoonhorst and colleagues (2011; see also Hoonhorst et al., 2009), compared categorical perception (i.e., phoneme discrimination) of French-speaking adults and children (aged from 6-8 years). They showed that relative to first-grade and pre-reading kindergarten children, older children who were also more advanced readers, were better at discriminating two sounds from the same phonetic feature continuum (i.e., /d/ and /t/ sounds). This improvement was indicated by the increase in the steepness of the identification slope (see Figure I.2). Kolinsky and colleagues (2021) have demonstrated recently that this increase in boundary precision (BP) is indeed due to the acquisition of literacy and is not driven by maturation and the related physical changes. With the aim to disentangle effects of maturation from those induced by reading acquisition, the authors tested beginning children readers, beginning adult readers, and skilled adult readers on the same categorical perception task employed by Hoonhorst and colleagues. Significant differences in BP observed between the two groups of adults, but not between children and adult beginning readers were interpreted as stemming from the acquisition of literacy (see Burnham, 2003 for more details on the *reading hypothesis*).

Figure I.2: Expample of an Identification function on a VOT continuum



Visual representation of how boundary precision changes after reading acquisition based on simulated data. The improvement in boundary precision, and consequently, categorical perception, is manifested as the increase in the steepness of the identification function (dark dotted line).

Finally, in one of our recent studies, differences in speech sound processing related to reading acquisition emerged (Jevtović, Stoehr, Klimovich-Gray, Antzaka, & Martin, 2022). We tested 60 second-grade Spanish-speaking children on perception and production of three speech sounds. Crucially, these sounds differed in the number of graphemes they map onto in Spanish. Consistent sounds /p/ and /t/ have only one orthographic representation in Spanish (i.e., <p> and <t> respectively), while the inconsistent sound /k/ is associated with three different graphemes (i.e., <c>, <qu> and <k>). Both speech sound perception and production of consistent sounds were modulated by children's reading skills. Better readers, who also developed stronger links between sounds and graphemes, were faster to perceive and produce consistent, but not inconsistent sounds.

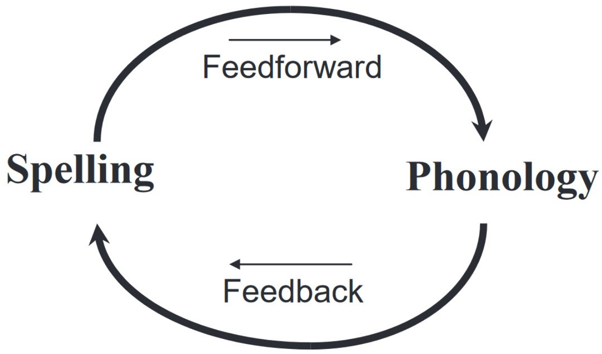
These studies reveal changes in speech sound processing related to reading acquisition. In the next section we present changes in word processing induced by reading acquisition, and specifically, the acquisition of sound-to-spelling correspondences.

The impact of orthography on spoken word processing

Another line of research showing the impact of orthography on speech processing comes from studies investigating how phonological and orthographic codes interact during visual word processing (see the previous paragraph). Previous research demonstrated that not only spelling-to-sound inconsistencies (i.e., orthographic representations with more than one possible pronunciation, hereafter *feedforward inconsistencies*), but phonology-to-spelling inconsistencies as well (i.e., phonological representations with more than one possible spelling, hereafter *feedback inconsistencies*; see Figure I.3)[2] hinder visual word recognition and reading (for a review, see Ziegler, Petrova, & Ferrand, 2008). Based on this finding, and in the light of interactive models proposing bi-directional links between different levels of representation (i.e., orthographic and phonological), researchers asked whether similar effects of feedback inconsistency can also be observed in the auditory domain. As a result, the studies in question made use of the fact that some words have consistent spellings (e.g., rhymes which can only be spelled in one way, such as the /ʌk/ in l**uck**), while others are inconsistent (i.e.., their rhymes have two or even more possible spellings such as /i:p/ l**eap** and j**eep**), to explore the potential effects of orthography on spoken language processing.

[2]Phonological and orthographic representations refer to both smaller units such as phonemes and graphemes, as well as larger units such as words. Therefore, an inconsistent sound is the one that maps onto multiple graphemes (e.g., the sound /k/ maps onto <c>, <qu> as well as <k>) and an inconsistent grapheme is the one that represents multiple sounds (e.g., the grapheme <c> represents both the sound /k/ as well as the sound /s/). Consistent sounds and graphemes by contrast, have one-to-one correspondences.

Figure I.3: Bidirectional Flow of Information Between Sound and Spelling



Schematic representation of the mutual influence of sound and spelling in language processing adapted from Van Orden and Goldinger, 1994.

Around the same time Morais and colleagues (1979) observed a link between phonological processing and reading skills, Seidenberg and Tanenhaus (1979) demonstrated that sound-to-spelling consistency affects spoken word processing. Interested in the possible interaction between orthographic and phonological codes during spoken language processing, the authors manipulated the spelling and the pronunciation of word pairs participants were presented with in a rhyme judgement task: some words overlapped in both their pronunciation as well as spelling (e.g., pie - tie), while the others shared the pronunciation but differed on the spelling (e.g., pie - eye). Faster decision times observed for word pairs containing rhymes with consistent spellings and pronunciations, were interpreted as the evidence of orthographic involvement during spoken language processing (see also Donnenwerth-Nolan, Tanenhaus, & Seidenberg, 1981 for similar results using better matched stimuli). The same effects of orthographic consistency have later been observed in *phoneme motioning* (Dijkstra, Roelofs, & Fieuws, 1995), *phoneme counting* (Treiman & Cassar, 1997) as well as *shadowing task* (Ventura, Morais, Pattamadilok, & Kolinsky, 2004). Importantly, observing the same spelling effects in paradigms in which orthography is *a priori* not relevant for the task, demonstrates that these spelling effects were not due to participants recurring to and explicitly relying on their orthographic knowledge. One such example is a lexical decision study in which Ziegler and Ferrand (1998) observed significantly longer response times for words containing rhymes which can be spelled in only one way (e.g., /oʊb/ in globe can only be written as <obe> in English) as compared to words whose rhymes have two or even more possible spellings (e.g., /eɪm/ in name can be written as either <ame> or <aim>).

The robustness and the generalizability of this *orthographic consistency effect (OCE)* are supported by studies showing that, apart from languages with highly inconsistent phonology-to-spelling mappings such as English and French (Ziegler & Ferrand, 1998; Ziegler et al., 2008), sound-to-spelling inconsistency also affects speech perception in more consistent languages such as Portuguese (Ventura et al., 2004), and Spanish (Jevtović et al., 2022). In addition, studies tracing the developmental trajectory of the OCE in normally-developing readers as well readers with dyslexia demonstrate that the effect is indeed driven by orthographic knowledge. The OCE seems to be absent in pre-reading (Ziegler & Muneaux, 2007) and children with dyslexia (Miller & Swick, 2003), who either have not yet acquired the orthographic knowledge, or have failed to acquire it successfully. Finally, although less robust (Alario, Perre, Castel, & Ziegler, 2007), evidence for OCE has been found in several speech production studies (Damian & Bowers, 2003; Rastle, McCormick, Bayliss, & Davis, 2011; but see Alario et al., 2007).

In summary, the aforedescribed effects show that once acquired, orthographic knowledge affects spoken language processing. Nevertheless, despite the large amount of evidence in favour of the OCE, the origin and the locus of this effect are still being debated in the literature. According to one line of research, for example, OCE is driven by the automatic co-activation of orthographic representations during spoken language processing (Chéreau, Gaskell, & Dumay, 2007; Pattamadilok, Perre, Dufau, & Ziegler, 2009). This means that upon hearing a spoken word, its orthographic representation is automatically activated leading either to facilitation (for words with unique spellings such as *name*) or slower recognition time (if the word has more than one spelling such as *globe*). By contrast, the *phonological* *restructuring account* suggests that with literacy acquisition, and in particular, the acquisition of sound-to-spelling conversion rules, the nature of the phonological representations changes (Muneaux & Ziegler, 2004; Perre, Pattamadilok, Montant, & Ziegler, 2009). Phonological representations of words with one possible spelling become more fine-grained and salient once reading is acquired, thereby leading to their faster recognition relative to spoken words with several possible spellings. Importantly, the two accounts are not mutually exclusive as it could be the case that both the automatic activation as well as phonological restructuring are driving the OCE.

The present thesis aims to add to the previous literature by exploring the role of orthography in *spoken word learning*. Spoken words' spellings have been shown to impact their auditory *perception* even when orthographic information is not relevant for the task. We thus set out to test whether orthographic knowledge can exert its effects during the acquisition of novel auditory word forms. Before presenting the objectives of the thesis, we will first discuss some previous literature investigating the link between orthography and novel word learning.

The role of orthography in word learning

Although the rate of learning novel words is not constant across the lifespan, as it tends to be the highest in the first years of life and during the schooling years (Anglin, Miller, & Wakefield, 1993), a vast amount of novel words are acquired all throughout adulthood (Brysbaert, Stevens, Mandera, & Keuleers, 2016; Nation & Waring, 1997). Despite its importance and omnipresence, the process of incorporating a novel word in the mental lexicon is not simple since it depends on the successful completion of several distinct steps. In particular, learning a novel word comprises acquiring a new phonological form (i.e., acquiring the knowledge about the sound of a particular word), acquiring the concept this novel word refers to (i.e., the semantics of a word), as well as learning the association between the two. In addition, the newly acquired word also needs to be integrated into the existing network of all already familiar words and the connections between them (McMurray, Kapnoula, & Gaskell, 2016).

A vast amount of psycholinguistic research has been dedicated to studying different factors that impact the abovedescribed learning process. For example, it has been shown that distributed learning (i.e., learning which is spread out over time) leads to better retention of novel word forms as compared to learning concentrated in one session (for a review, see Gerbier & Toppino, 2015). Similarly, experimental evidence shows that even though successful learning can occur incidentally (i.e., without the intention to learn), explicit instruction yields better learning outcomes (Batterink & Neville, 2011; Konopak et al., 1987; Sobczak & Gaskell, 2019). Finally, the importance of sleep for novel word acquisition and consolidation has repeatedly been shown in the literature (for a review, see Schimke, Angwin, Cheng, & Copland, 2021). Of particular importance for the present thesis however, are studies that looked into the role of orthography in novel vocabulary acquisition.

The idea that orthography may be involved in novel word acquisition, and can actually facilitate the learning process, comes from Ehri's word identity amalgamation theory (Ehri, 1978). According to this theory, orthographic labels (i.e., words' spellings) can be used as visual representations of phonological input, thus providing a way to symbolize (abstract) phonological representations when storing them in memory. Across four experiments Ehri and Wilce (1979) provided some of the first evidence supporting the ideas derived from the amalgamation theory. Namely, the authors reported a learning advantage for novel words acquired along with their orthographic representations - relative to those learned only via aural instruction - in both fist and second-grade English-speaking children. In addition, they described the conditions under which this advantage is most likely to occur. Firstly, they showed that spellings presented along with novel phonological forms have to represent the phonological input correctly, as misspellings can actually hinder the learning process. Secondly, and not surprisingly, they argued that in order to benefit from orthographic exposure, the children have to be able to decode the spelling correctly. They further demonstrated that the degree to which orthographic representations facilitate word learning positively correlates with children's reading skills. Finally, and of particular importance for the work presented in this thesis, they reported a similar learning advantage when children were only instructed to imagine words' spellings during the learning phase, but were not provided with any explicit orthographic information.

This link between orthographic information and word learning was further explored by Nelson, Balass and Perfetti (2005) who found that linking orthography to semantics is actually an easier task than linking semantic information to novel phonological input, thus supporting the ideas derived from Ehri's amalgamation theory (Ehri, 1978). Indeed, all studies reporting the positive contribution of orthographic information during word learning, denoted as *orthographic facilitation*, agree that orthography serves as an additional (visual) memory trace facilitating the learning process (Miles, Ehri, & Lauterbach, 2016; Ricketts, Bishop, & Nation, 2009; Rosenthal & Ehri, 2008).

Furthermore, orthographic facilitation was studied in a variety of populations, and was found to be significant not only in normally-developing children and adults, but children with hearing impairments (Salins, Leigh, Cupples, & Castles, 2021) as well as those with dyslexia (Baron et al., 2018). Despite their difficulties with phonological decoding, children with dyslexia seem to be able to benefit from orthographic exposure during word learning, at least when no verbal response is required in the task (Baron et al., 2018). Finally, orthographic facilitation effects, measured as both the number of correctly recalled novel words as well as the latency of the retrieval, were shown to hold even when learning novel words in a second language (Bürki, Welby, Clément, & Spinelli, 2019).

In sum, all the aforementioned studies show positive effects of orthography on learning novel spoken words. However, these studies explicitly used orthography in the learning process, thereby providing participants with an additional memorization cue. The present thesis investigates how orthography impacts word learning *indirectly*, that is, when it is not directly implicated in the learning process. Since novel words can be acquired without (explicit) reference to orthography (i.e., entirely in the auditory domain), it could be the case that, due to lifelong experience with written language, orthographic effects emerge even when spellings are not involved in the learning process. This hypothesis will directly be tested across three experiments conducted in the course of the present thesis.

The current thesis

The main goal of the present thesis is to explore the influence of orthographic knowledge on spoken word learning in adult skilled readers. Contrary to early childhood, when spoken language is the only source of vocabulary acquisition, encountering novel words as an adult can occur either aurally (in spoken language), through reading (in written language) or through spoken and written language at the same time. Therefore, we set out to explore how the two codes (i.e., phonological and orthographic) interact in the course of spoken word learning. As there is now vast evidence showing that *explicitly* linking orthographic labels to novel phonological words forms aids the memorization process (see above), the focus of the thesis was on the *implicit* effects of orthography on the learning process (i.e., when orthography is not explicitly present during learning). Specifically, the questions explored in this thesis were mainly motivated by two separate lines of research:

1) studies showing pervasive effects of orthography on spoken language processing (Seidenberg & Tanenhaus, 1979; Treiman & Cassar, 1997; Ziegler & Ferrand, 1998)

2) studies showing a positive link between children's oral vocabulary knowledge and their reading skills (Castles & Nation,

2008; McKague, Pratt, & Johnston, 2001; K. Nation & Snowling, 2004; Ouellette, 2006)

Since the studies showing orthographic effects in speech processing have already been discussed in detail, we will present the evidence in favour of the link between oral vocabulary knowledge and reading skills.

The relationship between oral vocabulary and reading

It is a well established finding that children's phonological skills, and specifically, phonological awareness, positively correlate and predict future reading success (Rack, Hulme, Snowling, & Wightman, 1994; Wagner & Torgesen, 1987). Researchers studying both normally-developing, as well as children with dyslexia, seem to agree that tasks tapping into phonological awareness skills are not only related to better reading skills (Swank & Catts, 1994) but can differentiate between good and poor readers even before the onset of reading acquisition (Turan & Guel, 2008; for a review, see Snowling & Hulme, 2021). With this in mind, it is not surprising that recent years have seen the emergence of studies focusing on less-studied predictors of reading success, one of them being oral vocabulary knowledge (K. Nation & Snowling, 2004).

Evidence for a link between oral vocabulary knowledge and reading performance is found in studies employing designs that vary in the degree of experimental control, and in consequence, the level of causality one can draw regarding the relationship between the two (Hulme & Snowling, 2013). For instance, positive correlations between vocabulary knowledge and word reading have been observed in several cross-sectional studies in which participants' standardized scores of oral vocabulary are used to predict reading success (Goff, Pratt, & Ong, 2005; Ouellette, 2006). Interestingly, the relationship between the two was found to be the strongest when reading words with irregular spellings (Ouellette & Beers, 2010; Ricketts, Nation, & Bishop, 2007).

More evidence for a link between oral vocabulary knowledge and reading performance is found when oral vocabulary knowledge is measured either at the early stages of reading acquisition, or even before its onset, and is used to predict future reading skills. In such a longitudinal study, which followed children from the age of 8.5 until 13, Nation and Snowling (2004) demonstrated that, in addition to phonological skills and listening comprehension skills, oral vocabulary knowledge accounted for unique variance in word reading scores. Contribution of orthography was found both the first time the testing was conducted, at the age of 8.5 years, as well as four and a half years later.

Finally, the most direct evidence for a link between oral vocabulary knowledge and reading success have been found in training studies using novel word material. One of the first studies to employ such a design was conducted with a group of fourth-grade English-speaking children. Children in this study were faster to read words previously acquired through aural training as compared to words they had never heard before, that is the untrained words (Hogaboam & Perfetti, 1978). Similar aural training advantage for reading previously acquired spoken words was found by McKague, Pratt, and Johnston (2001). The authors trained first-grade children on a set of novel spoken words and observed higher accuracy in reading words children had previously been trained on. However, while the two studies conducted with children showed that being familiar with a word's pronunciation is enough for the training advantage in reading to occur (i.e., semantic knowledge did not yield an additional benefit), in a training study conducted with adult readers, both semantic and phonological training were found to be important for reading success (Taylor, Plunkett, & Nation, 2011). As in these studies the amount and the type of training with novel words was experimentally manipulated, their conclusions thus point to a causal rather than an associative role of oral vocabulary in future reading success. However, the mechanism by which oral vocabulary assist word reading, is yet to be fully understood.

Theoretical explanations of the link between oral vocabulary and reading Although their main ideas and descriptions of how reading skills are acquired and developed differ considerably, most prominent theories of reading development agree that oral vocabulary assists word reading through a process underlying the acquisition of novel orthographic representations (i.e., *orthographic learning*; see Castles & Nation, 2008). For instance, according to Share's self teaching hypothesis (Share, 1995, 1999, 2004, 2008) already existing phonological representations (i.e., pronunciations of familiar words) serve as a reference point to which reading attempts of unfamiliar words are compared during the process of grapheme-to-phoneme conversion. Successful decoding attempts lead to a match between the written input and the existing phonological representation in the lexicon, therefore reinforcing the orthographic learning. In the same vein, unsuccessful attempts benefit from the phonological feedback (coming from the existing phonological representation) which then guide the reader to the correct their pronunciation. Similarly, Ehri's stage theory of reading development (Ehri, 2005, 2014) - which describes various stages of development children go through before becoming entirely skilled readers - postulates that, to form a novel orthographic representation, orthographic mapping between the spelling and the existing lexical representation, which contains both phonological and semantic information, has to occur. The beneficial role of oral vocabulary is seen as providing a semantic and phonological reference which guides the mapping or connecting process. Finally, according to the *lexical quality hypothesis* (Perfetti & Hart, 2001) phonological, orthographic as well as semantic information interact altogether in forming complete lexical representations of words in the mental lexicon. This then implies that better semantic and phonological representations should have a positive impact on orthographic learning (i.e., the process of generating a new orthographic representation).

Despite the important differences in how reading development is conceptualized - as a sequential process evolving through very specific stages (Ehri, 2005, 2014) or by acquiring orthographic representations at an item-level (Perfetti & Hart, 2001; Share, 1995) - all theories share the idea that the mechanism by which orthographic learning occurs operates *at the moment* of the first visual encounter with words' spellings. Namely, they aim to explain how letters (graphemes) get converted into sounds (phonemes) during first visual encounter with familiar words. An alternative explanation of the link between oral vocabulary knowledge and reading, which has been getting more attention in the recent years, proposes that the mechanism underlying the creation of orthographic representations can actually start *beforehand*. By converting sounds (phonemes) to letters (graphemes), a mechanism similar to the one employed when reading and sounding out unfamiliar words (Share, 1995), functions in a reverse manner (from sounds to letter) and generates preliminary orthographic representations of already familiar spoken words (see Stuart & Coltheart, 1988). As a result, a person's knowledge of phoneme-to-grapheme mappings can initiate orthographic learning even before the first visual exposure with the actual orthographic representations (i.e., words' spellings).

Some of the first experimental evidence showing that *orthographic expectations*[3] can be generated solely based on the novel phonological word form, comes from Johnston, McKague and Pratt (2004; see also McKague, Davis, Pratt, & Johnston, 2008). Across three experiments the authors demonstrated that English-speaking adults had encoded orthographic representations of novel words prior to the first visual encounter with words' real spellings. They first trained their participants on meanings and pronunciations of novel words without exposing them to words' spellings. Next, they presented participants with novel words' spellings in a modified lexical decision task containing a masked priming manipulation. Specifically, to tease out any potential effects of orthographic priming during access to novel visual words, they created four different

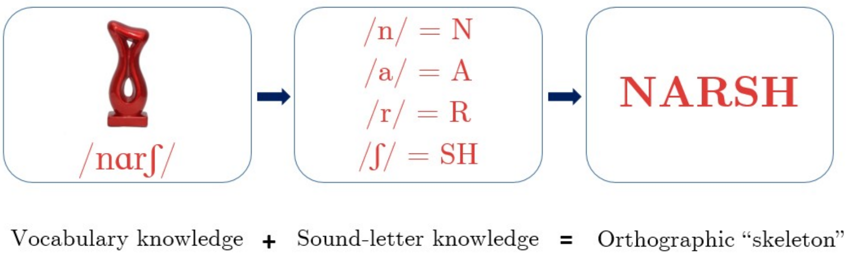
[3]The term orthographic expectations is used to denote the orthographic representation of a word that has been constructed before the first visual encounter with its written form, that is, solely based on phonological properties of the word.

prime-target conditions: pairs which overlapped in their phonological form (e.g., <spaith> before <SPATHE>), orthographic form (e.g., <spanth> before <SPATHE>), both (e.g., <spathe> before <SPATHE>) or neither (e.g., <gormin> before <SPATHE>). Since novel words preceded by identical primes (i.e., those overlapping in both the phonological and orthographic form) yielded faster processing times as compared to those preceded by primes with phonological overlap only, the authors could argue that the newly created representations of spoken words were not entirely phonological. In addition, as the masked priming task relies on automatic lexical access (Forster & Davis, 1984), their finding further demonstrated that representations of novel words are automatically accessed using the same recognition mechanisms employed when accessing already existing orthographic representations (i.e., representations of familiar words). Of particular importance for the authors' conclusions is the finding showing that novel words preceded by phonological primes and those preceded by purely orthographic primes yielded similar processing times. The latter were however processed significantly faster relative to novel words preceded by primes that were spelled using a completely different set of letters (e.g., <gormin> before <SPATHE>). The absence of purely phonologically mediated priming alongside the significant difference observed between orthographically same and different primes led the authors to conclude that these English adults had generated orthographic representations of novel spoken words during the learning phase. Notwithstanding, the authors themselves point out that participants could have imagined and generated spellings that differed from those they were later presented with. This leaves it unclear whether a single expectation was generated for each word or whether multiple expectations for alternative spelling patterns were considered.

The orthographic skeleton hypothesis

Recently there has been more evidence demonstrating that preliminary orthographic representations can be generated solely based on phonological word forms. This account has been known as the *orthographic skeleton hypothesis* (see Beyersmann et al., 2021; Wegener et al., 2018, 2020). According to this hypothesis, knowledge of phoneme-to-grapheme mappings can be used to generate preliminary orthographic representations (hereafter *orthographic skeletons*) for novel words acquired through aural instruction (see Figure I.4). Importantly, this happens even before the first visual encounter with novel words' spellings.

Figure I.4: Representation of the Orthographic Skeleton Hypothesis



Having a word in the oral vocabulary and knowing how sounds map onto letters will lead to the creation of a preliminary orthographic representation (i.e., the orthographic skeleton). For illustrative purposes we use American pronunciation of the example word.

The idea that oral vocabulary knowledge can assist word reading by generating orthographic skeletons was recently tested by Wegener and colleagues (Wegener et al., 2018). In their study, a group of English-speaking fourth-grade children was first trained on meanings (semantic training) and pronunciations (phonological training) of novel English-like words (e.g., /nesh/, /kɔɪb/, etc.). During the learning phase, which took place over two experimental sessions on two separate days, children were never exposed to orthographic input. Once the training was completed, children were presented with the novel words' spellings in a sentence reading task and their eye-movements were measured. Importantly, words were presented either in their predictable (e.g., /neʃ/ spelled as <nesh>) or unpredictable spellings (e.g., /kɔɪb/ spelled as <koyb> where <coib> would be more predictable). Along with the words children had previously been trained on (trained words), half of the sentences contained untrained words, again with predictable or unpredictable spellings. As the authors expected, and in line with previous research reporting a reading advantage for familiar spoken words (see above), there was a significant training effect. Trained words overall were read faster as compared to untrained words. The advantage was observed for all four eye-tracking measures: yielding shorter *total reading times*, shorter *gaze durations*, shorter *first fixation durations*, as well as fewer *regressions in* for trained words. Moreover, there was an overall facilitation for words shown in predictable spellings (e.g., <nesh>) as compared to those with unpredictable spellings (e.g., <koyb>). Finally, and crucially for the conclusions of the study, there was a significant interaction between training and spelling, again observed on all aforementioned measures. The interaction indicated that the facilitation for previously acquired words was present only when words were shown in their predictable spellings. Consequently, the authors interpreted this interaction between spelling predictability and training as evidence that children had generated orthographic expectations for all previously acquired words. Since the expectations children had generated for predictable trained words were in line with their real spellings, reading was facilitated. By contrast, the absence of facilitation for trained words with unpredictable spellings indicated a mismatch between orthographic skeletons children had generated for those words and the actual spellings they were later presented with.

These results were later replicated (using exactly the same procedure and materials) and expanded by showing that the effect of orthographic expectations persists even on the second encounter with the written forms of novel words (Wegener et al., 2020). However, after the third encounter, the facilitation effects disappear as children start updating their orthographic expectations based on their previous exposure with the real spellings. The authors of these two studies take these findings as a evidence that orthographic learning, that is, the process of acquiring novel orthographic representations, can start even before the first encounter with the written word forms. This is achieved through the creation of orthographic expectations (i.e., orthographic skeletons). Finally, using the same materials and word learning paradigm, it has been shown that the orthographic skeleton effects hold even when reading morphologically complex words containing stem participants acquired via aural instruction (Beyersmann et al., 2021). In addition, this study shows that adult readers (such as children tested until then) can also generate orthographic skeletons.

Along with Johnston et al. (2004), the authors of the aforementioned studies provide persuasive evidence for the orthographic skeleton account. They demonstrate that preliminary orthographic representations can be generated solely based on the words' phonological properties. The present thesis will further explore the process by which preliminary orthographic representations are generated in the course of spoken word learning by asking the following three questions:

1) Are orthographic skeletons generated for all novel words acquired through oral intruction or only for words with highly predictable spellings? By manipulating the number of novel words' possible spellings this question will be experimentally tested in Chapter 1.

2) Does generating orthographic skeletons depend on the complexity of phoneme-to-grapheme mappings (i.e., opacity of the language)? By comparing speakers of a language with simple, almost one-to-one mappings between sounds and letters (i.e., *transparent* languages such as Spanish) to those of a language in which this relationship is more complex since multiple sounds are associated to multiple letters, and vice versa (i.e., *opaque* languages such as French), this question will be experimentally tested in Chapter 2.

3) What is the nature of orthographic effects in spoken word learning? Specifically, are orthographic skeletons generated automatically during spoken word learning or, does generating them represents a conscious process participants strategically engage in to facilitate word learning? By comparing implicit versus explicit spoken word learning, this question will be experimentally tested in Chapter 3.

1 Learning spoken words with multiple possible spellings

Work presented in this chapter is based on:

Jevtović, M., Antzaka, A. and Martin, C.D. (2022). Gepo with a G, or Jepo with a J? Skilled Readers Generate Orthographic Expectations for Novel Spoken Words Even When Spelling is Uncertain. *Cognitive Science, 46*: e13118. https://doi.org/10.1111/cogs.13118

1.1 Introduction

The first experimental chapter of this thesis further explores the orthographic skeleton hypothesis by testing whether orthographic skeletons are generated for all novel words acquired via aural instruction, or only for words with a unique, and hence completely predictable spellings. Understanding the conditions under which orthographic skeletons are generated - or not - will shed light on how lenient the mechanism that drives phoneme-to-grapheme conversions is, and specifically, show whether it is constrained by spelling uncertainty. The idea, along with the supporting experimental evidence, that orthographic representations can be generated solely based on novel word's phonological form, have already been described in the introduction. Here, we will focus on questions that remained open, and which consequently motivated the present study.

Firstly, we saw that Wegener and colleagues (2018) manipulated the predictability of novel spoken words' spellings and showed a significant reading facilitation only for previously trained words with predictable spellings. The absence of facilitation for trained words with unpredictable spelling was interpreted as stemming from a mismatch between the correct spelling (i.e., the one children were presented with) and the one they generated during aural training (i.e., the orthographic skeleton generated during the learning phase). The authors thus concluded that orthographic representations for novel words can be generated solely based on the phonological properties of novel words, and importantly, that this occurs prior to readers' first visual encounter with words' actual spellings. Nevertheless, although their data provide strong evidence for the orthographic skeleton hypothesis, it remains unclear if the absence of the processing facilitation observed for words with unpredictable spellings, was caused by a mismatch between inaccurate orthographic representations children generated and the spellings they were presented with, as argued by the authors, or rather because children did not generate any representations for these words at all. Namely, it could have been the case that children did not even engage in the process of generating orthographic expectations when uncertainty regarding potential spellings was high. The first and the main goal of the present study was thus to test whether orthographic skeletons are generated even when there is uncertainty regarding the correct spelling (i.e., for words with more than one possible spelling), or only when words have a unique possible spelling. A way to adjudicate between these two possibilities, would be to train participants on a set of novel words controlled for the number of possible spellings. The idea being that, if a word has only one possible spelling, all participants would generate the same orthographic representation. If a word has multiple potential orthographic representations however, it is difficult to predict, on a participant level, whether orthographic representations would be generated, and if so, which of the possible and legal spelling option would be used to generate such representation. This leads us to the second goal of the present study. Namely, Johnston et al. (2004) pointed out that the spellings participants were presented with in their task might not have coincided with those they had previously generated. In particular, spelling expectations participants generated could have been different from those the experimenters provided in the task. However, if participants generate orthographic expectations even when multiple options are possible (i.e., under uncertainty), they should generate a one specific representation (selected from the possible options). This preferred spelling option, and the one likely to be used to generate the orthographic skeleton, could in some cases be based on statistical properties of languages (Davis & Perea, 2005). Nevertheless, as different spelling options can sometimes be equally possible (the case of the sound /b/ in Spanish; see below), it is difficult to predict which one would be used to generate a novel orthographic representation. Generating orthographic skeletons could thus, at least partly, be influenced by individual preferences. Therefore, as a second goal we wanted to see whether - if indeed generated - orthographic skeletons for words with more than one spelling follow individual spelling preferences.

Both of the aforementioned studies were conducted in English, an opaque language with extremely complex phoneme-to-grapheme conversion rules. Due to both irregular spellings and complex phoneme-to-grapheme mappings (English has 44 phonemes that can map onto more than 200 different graphemes), creating novel words with either a unique or only two possible spellings would be challenging, if not impossible. As a result, both predictable and unpredictable items are likely to have more than one possible spelling. Therefore, to control for the number of possible spellings and test whether orthographic skeletons are generated even under uncertainty, speakers of a more transparent language which allows us to better control for the number of possible spellings should be tested.

Contrary to English, Spanish orthography is made up of relatively simple phoneme-to-grapheme conversion rules (a total of 24 phonemes that map onto 32 graphemes). Crucially for the goal of the present study, most of the phonemes map onto only one grapheme (e.g., phonemes /m/ and /t/ can only be written as <m> and <t> respectively), and almost all vowels are completely consistent. This implies that there are many words (and consequently, pseudowords) whose spellings can be entirely predicted from their phonology, given that they have only one possible orthographic representation (e.g., the pseudoword /dalu/ can only be written as <dalu> in Spanish). At the same time, Spanish contains few inconsistent phonemes which have two orthographic representations (i.e., graphemes). The example that best illustrates Spanish orthographic inconsistency is the case of the sound /b/, which in Spanish maps either onto the letter <b>, or the letter <v>. Consequently, if only one sound in a particular word is inconsistent, that word would have exactly two (and not more) legal spellings (e.g., /bupe/ can be spelled as <bupe> or <vupe>). This property of the Spanish language provides a methodologically accurate way of controlling the predictability of novel word spellings: creating words with either only one or only two possible spellings. As as result, this enable testing whether orthographic skeletons are generated regardless of any uncertainty related to phoneme-to-grapheme mappings (both for words with a unique as well as those with those legal spellings) or are generated only when there is no uncertainty (i.e., only for words with a unique spelling such as /dalu/).

1.1.1 The present study

Experiment 1 aimed to further explore the orthographic skeleton hypothesis by testing whether orthographic representations are generated for all newly acquired spoken words or only for words with a unique and hence completely predictable spelling. To this end, a group of Spanish-speaking adults was trained on pronunciations of novel words with either only one (hereafter *consistent words*) or two possible spellings (hereafter *inconsistent words*). Namely, consistent words were made of phonemes with a unique possible grapheme representation in Spanish. As a result, spellings of these words were completely predictable from their phonology (e.g., /sufe/ can only be written as <sufe> in Spanish). By contrast, initial phoneme of all inconsistent words had two grapheme representations giving this way words two possible spellings (e.g., /χepo/ can be spelled either as <gepo> and <jepo>). Given that word spelling depends strongly on the graphotactic rules of a language (Carrillo & Alegría, 2014), for some words, the preferred spelling could be inferred and predicted based on bigram frequencies, as one of the two possible grapheme representations may be more likely to appear in a specific context. For instance, the phoneme /χ/ is more likely to be represented with the grapheme <g> when followed by the vowel /i/. For some Spanish phonemes however, it is more difficult to predict which grapheme is more probable to appear in a certain context given that the two representations have balanced frequencies (e.g., the sound /b/ followed by vowels /a/ or /o/ can be represented with either <b> or <v>, with similar frequencies in the language). This makes it difficult to anticipate, at the group level, which spelling would be preferred between the two options and consequently which would be used to generate an orthographic skeleton, if it is indeed generated despite uncertainty. Thus, there was a need to access participants' preferences for any of the two possible spellings. To account for possible individual differences in spelling preferences, we obtained obtained them beforehand. Namely, two weeks before the aural training took part, all participants provided their spelling preferences for all novel words (as well as some filler words; see Section 1.2.3.1). These preferences were used to present half of inconsistent words in each participant's preferred spellings and the other half in their unpreferred spellings. We assumed that if participants generate orthographic skeletons for aurally acquired words with two spellings, these orthographic skeletons would be based on individual preferences (i.e., the spelling option participants had provided beforehand).

Predictions regarding the outcomes of the study were the following. On the one hand, if participants generate orthographic expectations even when a word has more than one possible spelling, similar reading times should be observed for consistent and inconsistent words shown in their preferred (i.e., likely to be predicted) spelling. By contrast, words shown in their unpreferred spellings should elicit longer reading times indicating a mismatch between the expected and the real spelling. Given that this pattern should only be observed in the set of words participants had previously been trained on (i.e., trained words), a significant interaction between training and spelling should be found. As in Wegener et al. (2018), this interaction should be driven by a facilitation present when reading trained as compared to untrained words. Alternatively, the interaction could stem from significantly longer reading times present only for inconsistent unpreferred trained words (yielding a surprisal effect). On the other hand, if participants generate orthographic expectations only for words with one possible spelling, reading times should be faster for consistent (e.g., /dalu/) than for inconsistent (e.g., /χepo/) trained items regardless of the spelling (i.e., consistent words should be read faster as compared to either preferred or unpreferred inconsistent words). An interaction between spelling and training could be driven either by a facilitation present only for consistent trained words or by longer reading times observed for all inconsistent trained words (both preferred and unpreferred). Finally, no differences between the three spellings should be observed in the set of untrained words, given that in this case, no orthographic expectations could have been generated prior to the first visual encounter.

1.2 Methods

1.2.1 Participants

In order to have the equal number of participants across all experimental lists (see Section 1.2.3.3) as well as have a similar number of participants as were tested by Wegener et al. (2018), a sample size of 48 participants with usable data was determined before data collection.

A total of 54 participants completed the first session of the experiment. Due to technical issues three of them were not able to complete the second session and additional three participants had to be excluded from the analysis due to low accuracy in the aural training task. Data reported here come from the 48 participants (44 female) aged between 18 and 35 years (M = 25.6, SD = 3.74) who completed both experimental sessions within 14 to 16 days. All participants had Spanish as their first and dominant language, and their language skills were assessed through a series of objective proficiency measures: an interview conducted by a native Spanish speaker rated from 1 (lowest level) to 5 (native or native-like level), a picture naming task (the BEST proficiency test; de Bruin, Carreiras, & Duñabeitia, 2017), a lexical decision task (i.e., LexTALE-Esp which is the Spanish version of the LexTALE language proficiency test; Izura, Cuetos, & Brysbaert, 2014). In addition, subjective measures of proficiency were obtained through participants' self-reports on different aspects of proficiency. Namely, writing, listening, understanding, and speaking (see Table 1.1 for complete information about participants language profile). Participants were recruited from the BCBL Participa database and each received 15 euros for their participation in the study. The experiment was entirely web-based, but all participants had previous experience in participating in behavioral experiments in the laboratory and were therefore familiar with procedures and tasks used in experimental psychology (Uittenhove, Jeanneret, & Vergauwe, 2022). The experiment was approved by the BCBL Ethics Review Board (approval number 060420MK) and complied with the guidelines of the Helsinki Declaration. Participants' written consent was collected at the beginning of each experimental session.

Table 1.1: Participant Profile Experiment 1

	Mean	SD	Range
AoA	0	0	0-0
Picture naming (0-65)	64.7	0.54	63-65
LexTale (0-100%)	93	6.24	71.7-100
Interview (1-5)	5	0	5-5
Self-rated proficiency (0-10)			
Speaking	9.65	0.64	7-10
Understanding	9.64	0.61	8-10
Writing	9.45	0.77	7-10
Reading	9.51	0.75	7-10

Note. Some participants had some knowledge of a second or even a third language but none was highly proficient in any language other than Spanish.
[a] There are a total of 65 pictures to be named in the BEST (making 65 the maximum possible score).
[b] Self-rated proficiency data are missing for one participant.

1.2.2 Stimuli

1.2.2.1 Novel words Two sets of four-phoneme-long disyllabic novel words with the CVCV structure were created (set A and B, used as Trained and Untrained items and counterbalanced across participants). Each set consisted of eight consistent and 16 inconsistent words. Consistent words contained phonemes that map onto only one grapheme in Spanish. Consequently, consistent words had a unique possible spelling (e.g., /dalu/ can only be written as <dalu>). Inconsistent words by contrast, always stared with phonemes that had two possible orthographic representations in Spanish: /b/ which can be written as <b> or <v>; /k/ which when followed by vowels /i/ or /e/ can be written as <qu> or <k>; /ʎ/ which maps onto <ll> or <y>, and /χ/ which before vowels /i/ and /e/ has two orthographic representations <j> or <g>. This way, all inconsistent words had two legal spellings in Spanish (e.g., pseudoword /ʎedu/ can be written either as <lledu> or <yedu>). Half of the inconsistent words were a priori labeled as preferred and half as unpreferred. Preferred and unpreferred spellings were obtained for each participant individually through a pseudoword spelling task (see Section 1.2.3.1). Both preferred and unpreferred items contained two words starting with each of the four target inconsistent phonemes. Across preferred and unpreferred inconsistent items, words were also matched on the first syllable (two syllables per phoneme). Therefore, both preferred and unpreferred items started with the following eight syllables: /ʎe/ and /ʎu/, /χe/ and /χi/, /ba/ and /bu/, /ki/ and /ke/ (see Table 1.2).

Table 1.2: Novel Words from Experiment 1

Set	Consistent	Inconsistent Preferred	Inconsistent Unpreferred
A	/dalu/	/ʎedu/	/ʎefo/
	/duti/	/ʎuɲe/	/ʎupo/
	/femi/	/χepo/	/χede/
	/fipu/	/χifo/	/χitu/
	/ludi/	/bamu/	/badi/
	/nepo/	/bupe/	/bumi/
	/panu/	/kime/	/kifo/
	/muni/	/ketu/	/keli/
B	/dopu/	/ʎepo/	/ʎeli/
	/fadi/	/ʎule/	/ʎufi/
	/leme/	/χeni/	/χetu/
	/mepu/	/χipe/	/χidu/
	/nute/	/bafu/	/bani/
	/pimu/	/buɲe/	/buti/
	/sufe/	/kipe/	/kiɲo/
	/tamu/	/kefi/	/kedi/

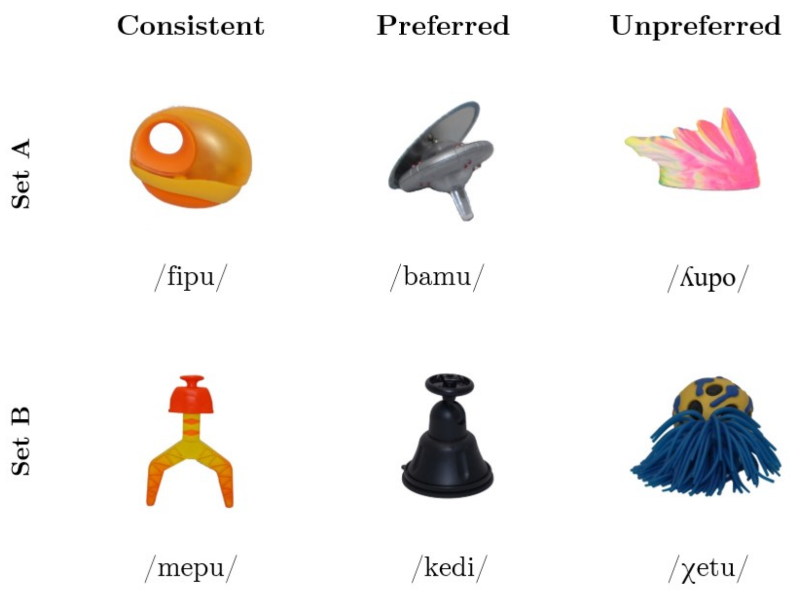
Note. Words from the inconsistent preferred group were later shown in each participant's preferred spelling whereas words from the inconsistent unpreferred group were presented in the unpreferred spelling.

Consistent items were matched between the two sets on the number of orthographic neighbours (no word had more than three neighbours, Set A: M = .750, SD = 1.04 and Set B: M = 1.38, SD = 1.22, $t(14)$ = -1.12, p = .281) and neither of the two possible spellings of inconsistent words had more than three orthographic neighbours. In order to avoid a potential gender mismatch between the demonstrative pronoun <este> preceding the novel in all experimental sentences (see Section 1.2.3.3) none of the novel words ended with the vowel /a/, as it most often marks the feminine gender of the nouns in Spanish (although there are some exceptions, e.g., $mano_{feminine}$). All items were recorded by a male L1 speaker of Spanish coming from the same region of Spain as the participants tested in the study. The recordings were made in a sound attenuated cabin using Marantz

PMD661.

1.2.2.2 Novel objects A total of 48 pictures selected from The Novel Object and Unusual Name (NOUN) Database (Horst & Hout, 2016) were used as the novel objects participants were presented with in the experiment (see Figure 1.1). Pictures were selected so as to be as different as possible from each other and were then randomly associated with novel words (see Table 1.2). Novel words and pictures associations were then kept constant for all participants (i.e., the same object was associated with the same word for all participants).

Figure 1.1: Novel Objects from Experiment 1



An example object from each set (set A and set B) and spelling group (Consistent, Preferred, and Unpreferred).

1.2.3 Procedure

The experiment was divided into two sessions, both completed online using the OSWeb online runtime, a JavaScript implementation of OpenSesame 3.3.2 software (Mathôt, Schreij, & Theeuwes, 2012) and hosted on the JATOS testing server (Lange, Kühn, & Filevich, 2015). All participants completed two experimental sessions with a two week pause in between them. Both sessions started with the audio message instructing participants to use headphones throughout the entire experiment as well as to make sure the sound is adjusted to a comfortable level.

During Session 1 participants first completed a pre-test spelling task, which aimed at assessing their individual spelling preferences. To mask the preferred spelling manipulation, during the same session participants completed two filler linguistic tasks: a lexical decision task followed by a real word spelling task. Tasks were always presented in this order so as to avoid any orthographic influence from the distractor tasks on the pre-test spelling task. Since these tasks were not of interest for the study, they will no be further discussed.

Session 2 always started with the aural training task during which participants were trained on the names of 24 novel objects. Next, they did a short distractor task (i.e., the Simon task; Simon, 1969) which was immediately followed by a self-paced reading task. In the self-paced reading task participants were for the first time exposed to the written forms of words they had previously been trained on. After the reading task, participants completed a short picture naming task in which they saw the pictures of all the objects they had been trained on and were instructed to name/spell them by typing the name of each object as they appeared one by one on the screen.

Finally, at the end of experiment participants were asked to answer several questions regarding the possible strategies they employed while doing the experiment, as well as hear their impressions about the tasks. They were also explicitly asked to indicate whether the spellings they were presented with differed from the ones they expected to see.

1.2.3.1 Pre-test spelling task The aim of the pre-test spelling task was to obtain participants' individual spelling preferences and determine the preferred and unpreferred spellings of inconsistent words for Session 2.

These preferences were then used to present words from the preferred group in spellings participants provided (i.e., preferred spellings), and those from the unpreferred group in the other, non-preferred spellings. For instance, if a participant spelled /bamu/, which belongs to the preferred group (see Table 1.2), with <b> in Session 1, that same participant would then see this word written with <b> in Session 2. By contrast, if a participant wrote /badi/, which belongs to the unpreferred group (see Table 1.2, with <b> in Session 1, that same participant would see the word written with <v> in Session 2.

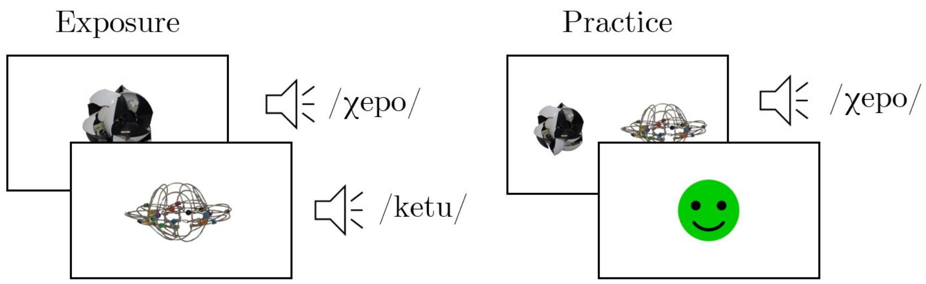
Participants were presented with a total of 96 novel words, half of them target (see Table 1.2) and half filler (see Table A.1). Filler words were created using the same phonemes and first syllables as the target words, and were added to ensure that participants would not remember, two weeks later during the aural training task, that they had already heard these novel words during Session 1. Participants were instructed to spell the novel words as if they were real words in Spanish (i.e., following the orthotactic rules of Spanish).

Novel words were presented in a randomized order and the task was self-paced. On each trial participants first heard a novel word over headphones and were prompted to spell what they heard in the text box appearing below the question “Please spell the word you just heard”. After typing in their response, they had to press ‘enter’ to move on to the next trial and hear the next novel word. The task took participants around 10 minutes to complete.

1.2.3.2 Aural training During the aural training task, participants were trained on the names of 24 novel objects taken from one the two sets (set A or set B which were counterbalanced across participants). They were instructed to learn the names of all objects since their performance would be tested later on in the experiment. In order to internalize the masculine gender of the nouns during the training phase, and thus avoid possible gender mismatch in the self-paced reading task (see Section 1.2.3.3), participants were explicitly told that the names of all objects were masculine.

With the aim to limit the learning load, novel objects and their corresponding names were presented in four blocks of six. Each block was made up of two words belonging to each of the three spelling groups (i.e., two consistent, two preferred and two unpreferred) and was split into two parts: exposure and practice part (see Figure 1.2). In the *exposure part*, participants were presented with the pictures of six novel objects. Pictures were presented one by one at the centre of the screen and at the same time the picture was presented, its name was played three times in a row at different speeds. The first and the third time, the name of the object was pronounced entirely, whereas the second time, it was pronounced by separating and emphasizing each of the two syllables (e.g., /muni/ -> /mu/ - /ni/ -> /muni/). After hearing the name of the object three times in a row, participants could press 'enter' in order to proceed to the next trial. Once they had been exposed to all six objects, they initiated the *practice part*. In the practice part, two different objects appeared on the left and on the right side of the screen, and at the same time, the name of one of them was played. Participants had to select the picture that corresponded to the name they heard by pressing 'M' or 'Z' on the keyboard (for picture on the right versus on the left). To reinforce the training process each trial was immediately followed by a feedback message (happy or sad face). Each picture was paired with every other picture and appeared once on the left and once on the right side of the screen, giving this way a total 60 trials in each practice block. At the end of each block, participants received a feedback message informing them about their overall accuracy rate (in %) in that practice block. This same procedure was repeated for all four blocks of words, and participants were encouraged to take breaks after each block.

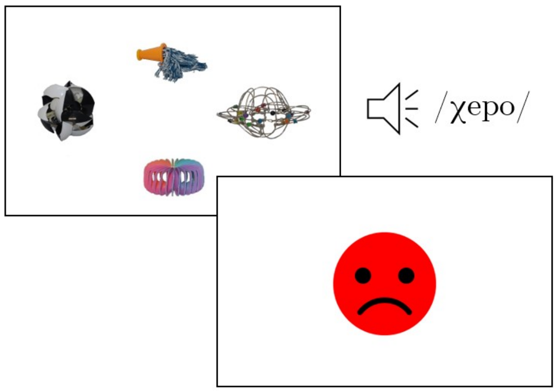
Figure 1.2: Trials of the Aural Training Blocks



In the exposure part (left), participants were presented with each of the six objects from that block one by one, while listening to their names spoken three times in a row. In the practice part, they saw two objects on the screen and heard the name of one of them. Participants had to select the object that corresponded to the name they had heard by pressing either 'M' (right) or 'Z' (left) on their keyboard.

Once all 24 objects had been presented and participants were trained to recognize their names, they completed the final block of the aural training task. In this block, participants were first presented with all 24 objects once again. Objects were presented one by one at the centre of the screen and at the same time, their names were played through headphones. As in the previous blocks, the exposure was self-paced, and participants moved from one picture to the next one by pressing 'enter' on their keyboard. After being familiarised once again with the names of all objects, they completed the final practice task. This time, four objects were presented on the screen, two on the left and the right side of the screen, and two on the upper and lower part of the screen (see Figure 1.3). As in the previous practice phases, at the same time the pictures were presented on the screen, participants heard the name of only one of them. To respond to which object on the screen corresponds the name they had heard, they had to use one of the four arrows on the keyboard (see Figure 1.3). To make sure that each picture appeared the same number of times at each of the four positions on the screen, and was paired equally with every other picture, position and pairing of the pictures was counterbalanced through Latin square, giving a total of 144 trials. As in the previous practice phases, participants received feedback message indicating whether their response was correct immediately after each trial. They also received the final feedback message informing them about their overall performance at the end of the task. On average participants needed approximately 30 minutes to complete the aural training task.

Figure 1.3: Final Block of the Aural Training Task



On each trial participants were presented with four objects on the screen and at the same time they heard the name of only one. After giving their response participants received a feedback message indicating whether the response was correct.

1.2.3.3 Self-paced reading task In the self-paced sentence reading task participants were presented with consistent, preferred and unpreferred spellings of the 24 words they had been trained on (i.e., trained words), as well as the 24 words from the other set (i.e., untrained words). Spellings of the words were embedded in eight different four-to-seven words long sentences (see Table 1.3). To make sure each target word appears the equal number of times in each sentence structure, and each of the six possible positions in the sentence (from two to seven), the position of the target word in a sentence was counterbalanced across the participants using the Latin square procedure. This yielded eight different sentence combinations repeated three times for the trained and three times for the untrained words (each participant read 48 sentences in total). Note that varying the length of the sentences, as well as the position of the target word within the sentence, was made to avoid the anticipation of the target word during reading (i.e., the moment of target word display within each sentence was unpredictable). Apart from making the task more engaging, the unpredictable position of the target word ensured that participants actually read the words, and do not just press 'enter' without processing them.

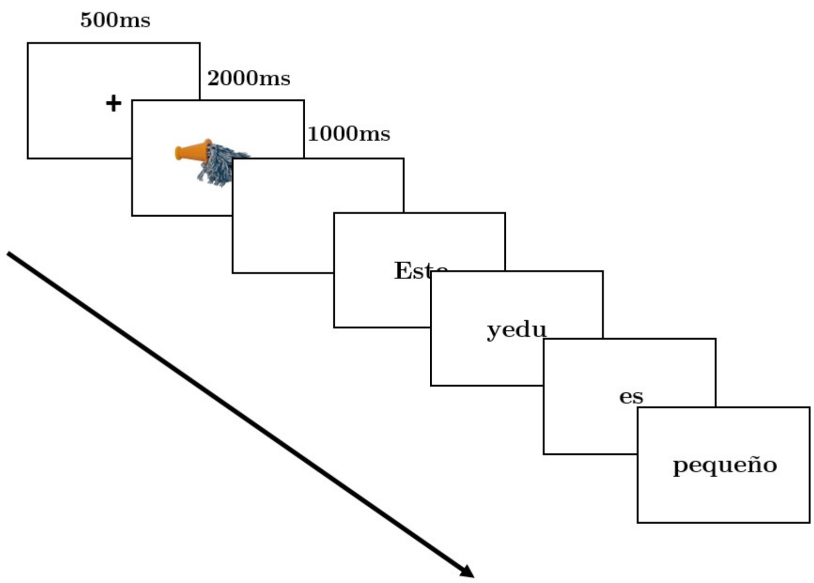
Table 1.3: Spanish Sentences used in the Self-Paced Reading Task

Original sentence	**English translation**
Este xxx es pequeño	This xxx is small
Este gran xxx es bonito	This big xxx is pretty
Este es un xxx grande	This is a large xxx
Este es un pequeño xxx fantástico	This is a small fantastic xxx
Este objeto es un xxx pequeño	This object is a small xxx
Este objeto es un pequeño xxx bonito	This object is one small fantastic xxx
Este gran objeto es un xxx fantástico	This big object is one fantastic xxx
Este gran objeto es un fantástico xxx	This large object is a fantastic xxx

Note. Due to syntactic differences across languages the position of the target word is not equivalent in the Spanish sentences and their English translations . Exes represent the place where target words appeared.

Sentences did not provide a rich semantic context but were preceded by a picture of the object whose named appeared in the sentence, this way priming the target word. Sentences were presented in a randomised order and the exact structure of each trial was the following: first a fixation cross appeared at the centre of the screen. After 500ms it was replaced by the picture of the object which stayed on the screen for a total of 2000ms. The picture was then replaced by a blank screen and after 1000ms the first word of the sentence appeared. The first word was the same across all sentences (i.e., each sentence started with the same demonstrative pronoun <Este>). From this point on, the task was self-paced and participants moved from one word to the next one by pressing 'enter' on their keyboard. After reading the last word of the sentence, participants initiated the beginning of the next trial by pressing 'enter' on their keyboard (See Figure 1.4). Participants were instructed to read each sentence as fast as possible without making pauses on any particular word. They were provided with six practice trials during which three sentences preceded by three familiar objects (e.g., a book, a glass and a pencil) were presented two times. Participants' reaction times, measured relative to the onset of the target word at the centre of the screen, were recorded and used as the index of their reading times. The entire task took participants around 10 minutes to complete.

Figure 1.4: The Structure of the Trial in the Self-Paced Reading Task



Each trial started with a fixation cross present for 500ms at the centre of the screen. The fixation cross was replaced by the picture of the object whose name was to appear in the following sentence. After 2000ms the picture of the object was replaced with a blank screen. The first word of the sentence appeared after 1000ms. Words appeared one-by-one at the center of the screen and the task was self-paced.

1.2.3.4 Picture Naming Task At the end of the experiment participants were presented with the pictures of the 24 objects they had been trained on, and were instructed to name them by spelling/typing their names. Pictures were presented in a randomised order one by one at the centre of the screen with the following trial structure: first, the picture of an object appeared at the centre of the screen. After 1000ms the picture became smaller, and a text box appeared below it, prompting participants to type in the name of the object shown on the picture. After writing the name of the object participants moved on to the next trial by pressing ‘enter’ on their keyboard. The aim of this task was to see how well participants could recall the names of the words they were trained on. The duration of the task was approximately five minutes to complete.

1.2.4 Data pre-processing and analysis

Data from the pre-test spelling task are presented in Appendix A.2 (see Table A.8 and A.9). Since they were only used to obtain individual spelling preferences, these data will not be further discussed.

To be included in the analysis of reading times (see Section 1.2.4.1) participants had to obtain at least 70% of accuracy in the aural training task. As data from the aural training task served only as an indicator of whether participants acquired the phonological forms of novel words they had been trained on, only descriptive statistics for the phonological training will be shown.

1.2.4.1 Self-paced reading task Reaction times for both trained and untrained words from the self-paced reading task were analyzed using linear mixed effects models (R. H. Baayen, Davidson, & Bates, 2008) in the statistical environment R (Version 4.2.0; R Core Team, 2021) and the R-packages *designr* (Version 0.1.12; Rabe, Kliegl, & Schad, 2021), *ggplot2* (Version 3.3.6; Wickham, 2016), *lme4* (Version 1.1.29; Bates, Mächler, Bolker, & Walker, 2015), and *lmertest* (Kuznetsova, Brockhoff, & Christensen, 2017) were used for data analysis and visualisation.

Before analyzing reaction times (RTs), extreme values (i.e., RTs below 150ms and above 1200ms, representing 4.9% of the data) identified based on the visual inspection of the raw data were removed (Ratcliff, 1993; see also R. H. Baayen & Milin, 2010). [4] In line with the Box-Cox test (Box & Cox, 1964) RTs were then log transformed in order to improve the positively skewed distribution, as well as minimize the effects of any possible outliers

[4]Note that the lower rate of outlier removal (e.g., removal lower than 1.5 % of the data or even no removal at all) yielded the same pattern of significance in the main analysis.

(R. Baayen, 2008).

To test whether there was a significant difference in reading words with a unique spelling (i.e., consistent words) and each of the two inconsistent spellings (i.e., inconsistent preferred and inconsistent unpreferred spellings), the three level factor group spelling was deviation coded (Schad, Vasishth, Hohenstein, & Kliegl, 2020). The first contrast thus compared the RTs between consistent and inconsistent preferred spellings (hereafter *Spelling1vs2*; consistent spellings were coded as -0.33, inconsistent preferred as 0.67 and inconsistent unpreferred as -0.33) while the second contrast compared consistent to inconsistent unpreferred spellings (hereafter *Spelling1vs3*; consistent spellings were coded as -0.33, inconsistent preferred as -0.33 and inconsistent unpreferred as 0.67). The fixed factor *Training* was initially deviation coded (trained words were coded as 0.5 and untrained as -0.5). In case of a significant interaction between training and either of the two contrasts, training was treatment coded (i.e., the level of interest was coded as 0 and the other as 1) in order to change intercept and thus look at the simple effects of the contrast either at the level of trained words only (trained as 0 untrained as 1) or untrained words only (trained as 1 untrained as 0). In addition, the two-level factor *Set* representing to two sets of items, was deviation coded (Set A as -0.5 and Set B as 0.5) and included in the model as the fixed covariate.

To avoid overfitting the statistical models which would lead to reduced statistical power (Matuschek, Kliegl, Vasishth, Baayen, & Bates, 2017), random-effects structure was build following a parsimonious data-driven approach (Bates, Kliegl, Vasishth, & Baayen, 2015). As a result, all reported models represent the highest nonsingular converging models and include the maximal random effects structure for all experimental manipulations of interest supported by the data (see Table 1.5 for exact model structure).

1.2.4.2 Picture naming task In the picture naming task, participants had to spell the names of the 24 objects they had been trained on. Given the complexity of the task (i.e., spelling the entire word correctly) to derive a more precise measure of word recall accuracy (regardless of the spelling preference) instead of considering accuracy as a binary 1 (correctly spelled word) or 0 response (incorrectly spelled word), phonetic distance between the correct and given response was calculated for each response using the ALINE string alignment algorithm. The ALINE algorithm quantifies the phonetic similarity between any two written strings. This measure is considered complementary to Levenshtein distance, as apart from the number of insertions, deletions and substitutions, it also takes into account phonetic features when calculating the similarity between two strings (Kondrak, 1999).

The ALINE distance for each response and each participant was calculated using the *alineR* package in R (Downey, Sun, & Norquest, 2017). The *alineR* package relies on the ALINE algorithm to calculate the phonetic similarity score between any two word strings and gives a value from 0 (no difference) to 1 (completely different strings). Given that the phonetic similarity score between the two identical strings results in 0 (no differences), meaning that higher values indicate lower accuracy, to ease the interpretation of the score, the Inverse ALINE score, which was calculated as 1-ALINE (i.e., higher values stand for higher accuracy and better recall), is reported here. Note that, since 1-ALINE is indexing phonetic (rather than orthographic) closeness, the two plausible spellings of an inconsistent word would both give a score of 1 (i.e., both <bumi> and <vumi> for the word /bumi/ give the same score).

Before the analysis, responses where participants did not write anything ('empty responses') as well as responses representing real Spanish words were removed (4.51% of all data).

1.3 Results

1.3.1 Aural training

In the aural training task, participants were thought the names of 24 novel objects. Overall accuracy in the final check phase was high: 91.4% (*SD* = 8.01, range 70%-100%), compared to an at chance level of 25%. Importantly there were no significant differences in accuracy between the two sets of words (Set A: M = 92.1, SD = 8.62; Set B: M = 90.6, SD = 7.45; $t(46)$ = 0.627, p = .534). Three participants who obtained less than 70% of accuracy were excluded from any further analysis. Moreover, accuracy per training block (see Table 1.4) was also high (>90% in all blocks), with an at chance level of 50%.

Table 1.4: Accuracy in the Spanish Aural Training Task

	Block1	**Block2**	**Block3**	**Block4**	**Final Check**
Set A	93.6 (4.61)	96.2 (4.77)	97.7 (2.61)	94.9 (5.09)	92.1 (8.62)
Set B	95.0 (4.91)	97.7 (2.72)	96.2 (4.20)	95.3 (4.49)	90.6 (7.45)

Note. Mean percentage of accuracy (SDs) per training block and in the final block of the aural trainig task.

1.3.2 Self-paced reading

In the self-paced reading task, participants were presented with the spellings of aurally acquired novel words, as well as with words from the untrained set. Both trained and untrained words were presented in their unique (i.e., consistent words), preferred or unpreferred spellings (i.e., inconsistent words). Mean RTs for all target words, measured relative to the onset of the word at the center of the screen, are shown in Figure 1.5. Their distributions in the form of a raincloud plot are shown in Figure A3 in Appendix A.3).

The full model structure including both fixed and random effects is shown in Table 1.5. The main model looking into all words (i.e., both trained and untrained) showed no main effects of either the *Spelling1vs2* (β

= .013, *SE* = .017, *t* = .741, *p* = .458) or the *Spelling1vs3* contrast (β = .034, *SE* = .018, *t* = -.059, *p* = .064). The main effect of *Training* was significant (β = -.038, *SE* = .017, *t* = -2.25, *p* = .029) indicating that untrained words overall (*M* = 419, *SD* = 144) were read faster as compared to the trained words (*M* = 432, *SD* = 140). Importantly, while the interaction between training and *Spelling1vs2* contrast was not significant (β = .014, *SE* = 0.034, *t* = .394, *p* = .693) the *Spelling1vs3* contrast interacted significantly with training (β = -.069, *SE* = 0.034, *t* = -2.02, *p* = .043).

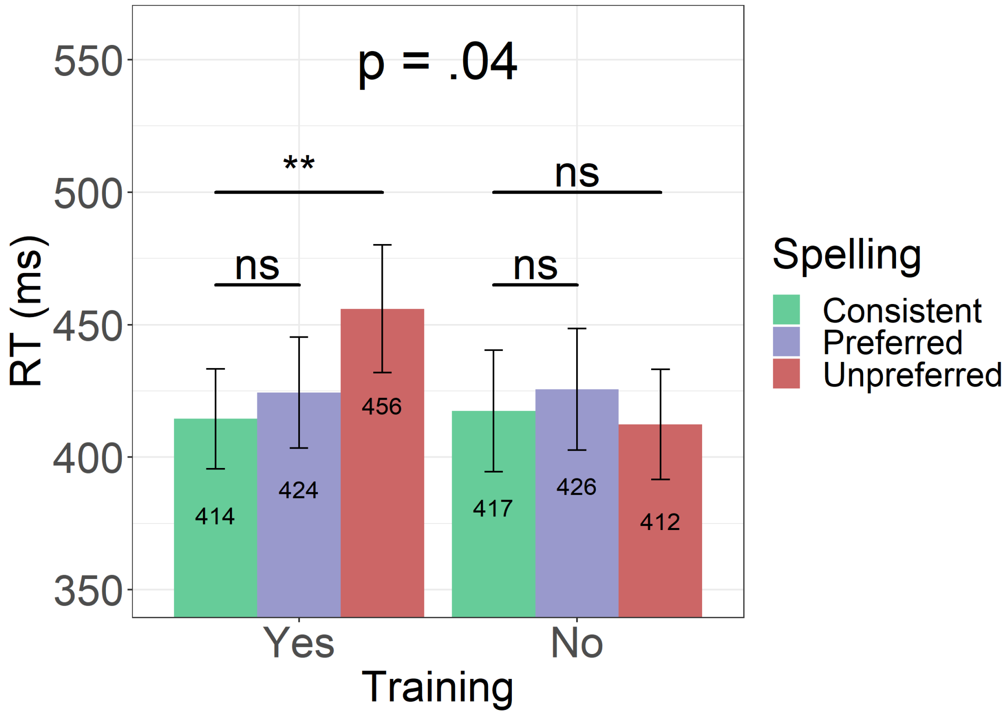
Table 1.5: Full Structure of the Model from Experiment 1

Fixed effects	**Estimate**	**SE**	***t* value**	***p***
(Intercept)	5.94	0.048	124	**< .001*** **
Training	-0.038	0.017	-2.25	**< .05***
Spelling1vs2	0.013	0.017	0.741	0.458
Spelling1vs3	0.034	0.018	1.88	0.064
Set	-0.140	0.096	-1.46	0.152
Training: Spelling1vs2	0.014	0.034	0.394	0.693
Training: Spelling1vs3	-0.069	0.034	-2.02	**< .05***
Random effects	**Variance**		**Std. Dev.**	
Participant (Intercept)	0.108		0.328	
Participant: Training (slope)	0.004		0.065	
Participant: Spelling1vs3 (slope)	0.001		0.038	

Note. The full model looks into both trained and untrained words.

Further inspection of the interaction between *Training* and *Spelling1vs3* contrast, performed by treatment coding the factor *Training* and hence changing the reference level (first to trained and then to untrained words only), showed that, while the difference between consistent and unpreferred spellings was significant in the group of trained words (β = .069, *SE* = .025, *t* = 2.76, *p* = .006), the same difference failed to reach the level of significance when only untrained words were considered (β = -.001, *SE* = .025, *t* = -.032, *p* = .974).

Figure 1.5: Reaction Times From the Experiment 1



Reaction times for for both trained (Yes) and untrained (No) consistent, preferred and unpreferred word spellings. Error bars represent the standard error of the mean, and the numbers within the bars represent the menans for each condition.

To sum up, while no differences in RTs were found between consistent and either of the two inconsistent spellings (preferred or unpreferred) in the group of untrained words, trained words presented in their unpreferred spellings yielded significantly longer reading times as compared to the words with consistent spellings.

1.3.2.1 The role of bigram frequencies To explore the role of bigram frequencies in generating orthographic skeletons, instead of considering participants' individual preferences, inconsistent words were split in two groups (i.e., classified as preferred or unpreferred) based on bigram frequencies of the target syllables. Bigram frequencies were calculated using B-Pal (Davis & Perea, 2005), and words with initial syllables <ba>, <bu>, <je>, <gi>, <lle>, <llu>, <que> and <qui> were classified as preferred, since they are more frequently present in the initial position of real Spanish words. Their counterparts (i.e., less frequent initial bigrams) were hence considered as unpreferred.[5]

The full fixed-effects and random-effects structure of the model looking into bigram frequencies is shown in Table 1.6. The model showed neither the interaction between *Training* and *Spelling1vs2* contrast (β = -.018, *SE* = .035, t = -.524, p = .600) nor the interaction between *Training* and *Spelling1vs3* (β = -.043, *SE* = .034, t = -1.26, p = .206).

Therefore, the pattern of results observed when considering participants' personal preferences was not replicated when preferred and unpreferred spellings were inferred from statistical properties of the language (i.e., the bigram frequency).

Table 1.6: Full Structure of the Model Looking into Bigram Frequencies

Fixed effects	**Estimate**	**SE**	***t* value**	***p***
(Intercept)	5.94	0.048	124	**< .001*** **
Training	-0.039	0.017	-2.31	**< .05***
Spelling1vs2	0.031	0.021	1.50	0.140
Spelling1vs3	0.034	0.017	1.96	0.054
Set	-0.139	0.096	-1.44	0.154
Training: Spelling1vs2	-0.018	0.035	-0.524	0.600
Training: Spelling1vs3	-0.043	0.034	-1.26	0.201
Random effects	**Variance**		**Std. Dev.**	
Participant (Intercept)	0.108		0.329	
Participant: Training (slope)	0.004		0.066	
Participant: Spelling1vs2 (slope)	0.007		0.081	
Participant: Spelling1vs3 (slope)	0.001		0.024	

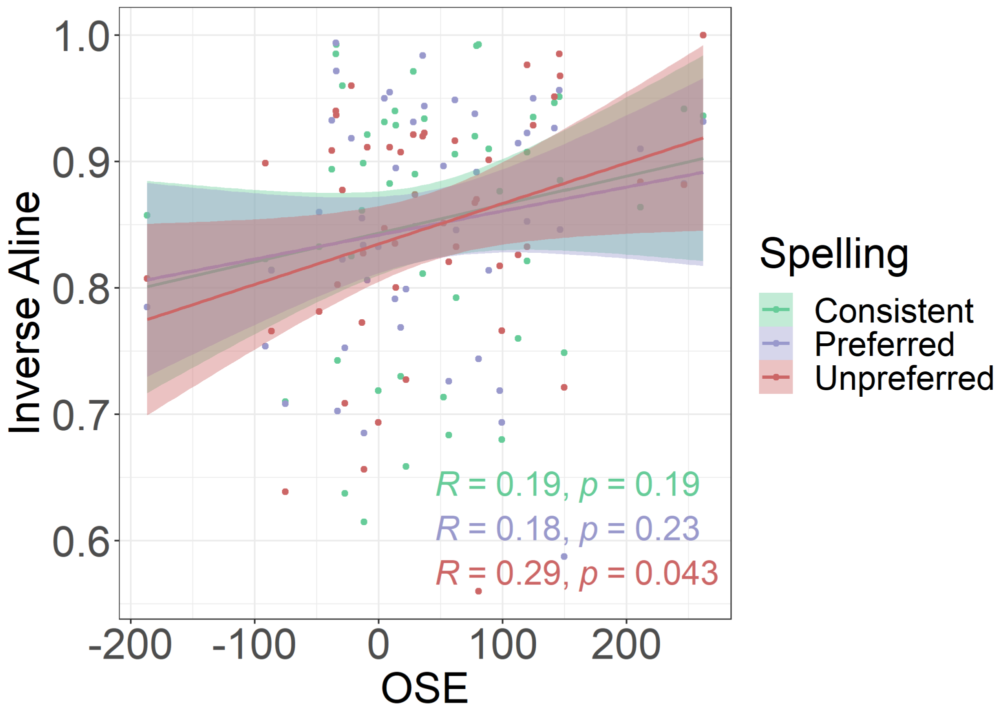
Note. The full model looks into both trained and untrained words.

[5] Note that this led to an uneven number of items across the two inconsistent groups of items. Namely, given that the items that were for some participants classified as preferred, for the others, those same items were classified as unpreferred. For instance, the item /bumi/, which was originally a preferred item, had to be presented to some participants as <bumi> and to some as <vumi> in the reading task (based on individual spelling preferences). By contrast, following only the bigram frequency, this item was classified as preferred for all participants who saw it with the letter <b> in the task, and as unpreferred for those who saw it with <v>.

1.3.3 Exploratory analysis

As an exploratory question we were interested in the potential relationship between generating orthographic skeletons and retaining the newly acquired words. A measure of participants' tendency to generated orthographic skeletons was operationalised as a difference in reading previously acquired words shown in their unpreferred and those with consistent spellings (i.e., mean RT) and denoted as the *orthographic skeleton effect (OSE)*. Each participant's OSE score was then correlated with their performance on the picture naming task (i.e., their inverse ALINE score). Correlations were run for the three different spelling groups separately (see Figure 1.6).

Figure 1.6: Correlation Between the OSE and Word Recall



The OSE (on the x-axis) operationalized as a difference in mean reaction times between the unpreferred and consistent spellings. The higher the value, the larger the effect. Analogously, higher inverse ALINE distance (y-axis) represents higher accuracy in recalling the names of the objects participants were trained on.

As shown in the Figure 1.6, the only significant correlation was the one found between the OSE and the inverse ALINE score for words from the unpreferred spelling group ($r(46) = .29$, $p = .043$). This correlation indicates that participants who were more susceptible to generate orthographic skeletons as a result of aural training, were the ones who remembered better the novel words presented in unpreferred spellings in the reading task.

1.4 Discussion

The first experiment of the present thesis aimed to further explore the orthographic skeleton hypothesis by testing whether preliminary orthographic representations are generated for all newly acquired spoken words, or only for words with a unique and hence entirely predictable spellings. To that end, 48 adult speakers of Spanish participated in a two-session online experiment. During Session 1, participants' individual spelling preferences were collected for all novel words (i.e., both trained and untrained) through a pre-test spelling task. Two weeks later, during Session 2, participants were trained on the pronunciations of novel words that varied in terms of the number of possible spellings: consistent words had only one, while inconsistent words had two possible spellings. Following the aural training, participants were presented with spellings of both trained and untrained words in a self-paced sentence reading task. Significantly longer reading times were observed only for newly acquired words presented in their unpreferred spellings, thereby suggesting that Spanish adult readers had generated orthographic expectations as a result of phonological training. Importantly, they did so for both consistent and inconsistent words (i.e., when there was certainty regarding spelling, but also when there was uncertainty due to inconsistent phonemes). Furthermore, orthographic skeletons for inconsistent items were in line with participants' individual spelling preferences, given that inconsistent words shown in their preferred spellings did not differ in reading times from words with a single possible spelling (i.e., consistent words). As no significant differences were observed between consistent and inconsistent untrained words, this way yielding a significant interaction between spelling and training, we interpret these findings as evidence that participants indeed generated orthographic skeletons during aural training.

The present study introduces two important innovations: firstly, the number of possible spelling options for each novel word was controlled for by creating words with only one, or two possible spellings. Secondly, participants' personal spelling preferences were used to determine the preferred spelling option for words with two possible spellings. As a result, we were able to overcome a caveat present in the study conducted by Wegener et al (Wegener et al., 2018) related to the stimuli they used. Since due to high complexity of phoneme-to-grapheme mappings present in the English language, items with predictable and unpredictable spellings could not be matched on number of graphemes, as well as bigram frequency, observed differences between these two groups of items could be at least partly linked to stimuli properties rather than orthographic skeletons participants generated during the learning phase.

Additionally, by assessing individual spelling preferences, and this way determining the orthographic skeleton participants were more likely to generate when two options were available, the present study was also able to address an issue raised by Johnston et al., (2004) concerning individual variability in orthographic expectations. Indeed, in the present study preferred spellings varied considerably across participants (see Appendix A.2). Interestingly, in some cases, preferred spellings deviated from the orthotactic rules of the language. For example, based exclusively on the frequency of its appearance in Spanish, the grapheme <ll> should be preferred for items with the /ʎ/ sound at the initial position. Looking at Tables A.8 and A.9, it is clear that this was not the case, since the majority of the participants tested in the study actually preferred the less frequent grapheme <y>. More importantly, as indicated by the absence of an interaction between spelling and training when bigram frequencies, rather than individual spelling preferences, were considered, the latter were indeed favored in generating orthographic skeletons for novel words with two spellings. Nevertheless, although this manipulation allowed us to adapt the stimuli material at the participant level, it introduced a potential confound. Namely, it could be that case that obtaining participants' spelling preferences beforehand may have influenced their performance on tasks completed two weeks later. To minimize the impact of the pre-test spelling several precautions had been taken: additional filler tasks and filler items were added, and a two-week delay between the sessions was introduced. These precautions seem to have been enough to mask any potential influence of the pre-test task, given that significant differences in reading times were present only for previously trained items, despite the fact that both trained and untrained items have been presented in Session 1. In conclusion, future studies dealing with novel word spellings could, in addition to considering other important psycholinguistic variables, also adapt their material at the participant level.

Further evidence for the orthographic skeleton account comes from the exploratory analysis in which participants' tendency to generate orthographic skeletons (i.e., the OSE score calculated as a difference in reading times for unpreferred and unique spellings) was correlated with their performance on the picture naming task at the end of the experiment. The significant positive correlation observed only between the OSE score and the accuracy in remembering words shown in their unpreferred spellings, suggests that participants who are more prone to generating orthographic skeletons are also more likely to correctly recall the items encountered in the unexpected spelling, that is, those that deviated from their expectations. The observed relationship between word recall and the surprisal effect in reading (i.e., longer reading times for unpreferred spellings), should however be explored in more detail in order to draw any strong conclusion about its nature. For instance, participants' tendency to generate orthographic skeletons should be investigated in relation to other constructs shown to play a role in word learning (e.g., phonological short-term memory).

1.4.1 Comparison with Wegener et al. (2018; 2020)

The findings from the present study add to the previous literature by showing that the mechanism driving phoneme-to-grapheme conversions functions even under uncertainty. They therefore expand on the findings reported by Wegener and colleagues (2018, 2020). Nevertheless, this and the two studies by Wegener and colleagues differ in their methodological approaches, leading to differences in the observed results. Firstly, the interaction between spelling and training observed in the present study, was driven by longer reading times present only for previously acquired inconsistent words shown in their unpreferred spellings. The interaction observed in the study conducted by Wegener et al. (2018, 2020) however, stemmed from shorter reading times and consequently a significant facilitation observed only for previously trained words shown in their predictable spellings. We propose several explanations for this reversed pattern of results, and in particular, the absence of training advantage in the present study: first, novel word learning paradigm employed by Wegener et al. (2018, 2020), apart from being more extensive and distributed over two experimental sessions, included both aural and semantic training. By contrast, participants in the present study were exposed to relatively short aural training, with only the picture of the object as semantic context. Their orthographic expectations were tested immediately after, in the course of the same experimental session. This short training, as well as the delay between training and testing, might not have been long enough for the effect of aural training to emerge. Secondly, different techniques were employed, and consequently, the evidence for the orthographic skeleton hypothesis is based on different dependent measures. While the conclusions of the present study come from reading latencies measured through a behavioral response (i.e., button press after reading a word), Wegener and colleagues employed an online measure of the reading process (i.e., eye-tracking). In the same vein, the simplicity of the reading task employed in the present study, which consisted of reading short disyllabic words embedded in relatively simple sentences with no semantic context (see Section 1.2.3.3), might have compromised the likelihood of detecting the training facilitation observed in the previous aural training studies (e.g., Álvarez-Cañizo, Suárez-Coalla, & Cuetos, 2019; Michael Johnston et al., 2004; McKague et al., 2001). Moreover, these studies were conducted in languages with highly distinct writing systems. English, which has been used in all previous studies, has a highly inconsistent orthographic system with both phoneme-to-grapheme as well as grapheme-to-phoneme inconsistencies. This means that both reading as well as spelling unfamiliar words in English comes with high uncertainty (Ziegler, Stone, & Jacobs, 1997). As a result, English speakers are frequently confronted with unexpected spellings. Their scarce experience with predictable spellings could have lead to the facilitation effects in reading previously acquired spoken words in line with their expectations. By contrast, Spanish speakers are rarely confronted with irregular spellings, and may hence be more sensitive to situations in which their expectations are not confirmed (as indicated by longer reading times only for unpreferred spellings). Apart from different techniques, different paradigms and designs, differences in the writing systems also partly explain why trained words did not lead to an overall processing advantage in the current study. Finally, we cannot discard the possibility that differences in reading skills could have also lead to differences in reading trained versus untrained words. In the present study adult skilled readers were tested, whereas Wegener and colleagues tested children developing readers. Compared to developing readers who rely heavily on their phonological decoding skills when reading, and thus need more time to sound out unfamiliar words (Share, 1995), skilled readers are due to their extensive experience and automaticity in reading fast even when reading unfamiliar words. Therefore, any differences between trained and untrained words present in skilled readers might haven been too small to be detected in reading times measured via button presses.

1.4.2 Conclusion

Before concluding, one important limitation of the present study, which will be explored in the remainder of the present thesis, should be mentioned. The conclusion that orthographic expectations are generated even when there is uncertainty regarding the possible spelling (i.e., even for inconsistent items) is based on reading novel words with only two possible spellings. However, it is possible that orthographic skeletons are not generated when the number of possible word spellings larger, and consequently the probability of a mismatch between the expectations and the actual spelling is higher. Moreover, the process of generating orthographic skeletons might also be constrained by the complexity of the specific orthographic system. In languages like Spanish, the overall probability of generating an incorrect orthographic representation is low. As a result, Spanish speakers may be more prone to generating orthographic expectations even when there is uncertainty regarding the correct spelling. Therefore, the ubiquity of skeleton creation when learning novel words with multiple spellings should be explored in other languages with more opaque writing systems. Due to more complex sound-to-spelling rules, opaque languages contain higher level of uncertainty, and consequently higher risk of error when generating orthographic skeletons than Spanish. The generalisability of the results we report here will be investigated in Experiment 2.

2 The role of orthographic depth in spoken word learning

Work presented in this chapter is based on:

Jevtović, M., Antzaka, A. and Martin, C.D. (in press). Déjà-lu: When orthographic representations are generated in the absence of orthography. *Journal of Cognition*

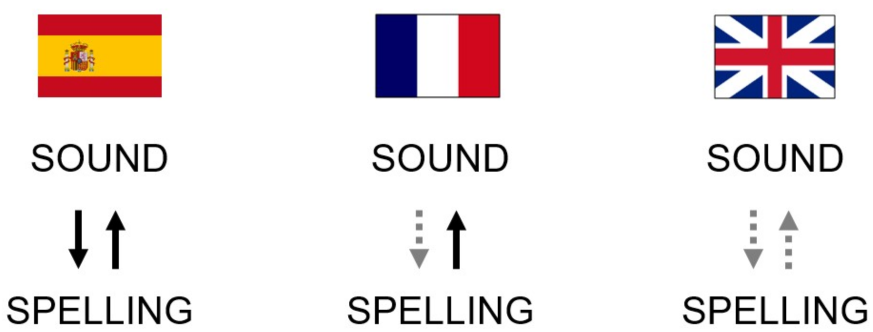
2.1 Introduction

In the previous chapter we obtained further evidence for the orthographic skeleton account by showing that orthographic skeletons are generated even for words with two possible spellings. This finding suggests that the mechanism driving phoneme-to-grapheme mappings, which in return lead to the creation of orthographic skeletons, functions even under uncertainty caused by two possible spelling options. In addition, we showed that participants tend to follow their own spelling preferences, rather than statistical properties of the language, when generating orthographic representations for newly acquired spoken words. However, to test the leniency of this process, and see whether orthographic skeletons are generated even when there is uncertainty regarding novel words' spellings, we had to control for the number of possible spellings. This was done by creating words with unique, or only two possible spellings. Given that suchlike manipulation was not feasible in a highly inconsistent language like English, we tested adult speakers of Spanish. Consequently, the obtained results may very much be limited to the speakers of transparent languages. Indeed, Spanish is a language in which, in most cases, the predicted spelling matches the real one. Due to the overall low risk of error, speakers of transparent languages could be more prone to generate orthographic representations even when more than one spelling is possible. By contrast, speakers of opaque languages such as English, may refrain from generating orthographic skeletons since the probability of generating a correct one (i.e., the one that will match the real spelling) is lower. In the second experiment we expand on these findings by testing how generalisable they are. In particular, we set out to investigate whether generating orthographic skeletons for words with two spellings occurs even when the *overall* probability of generating an incorrect representation is high, as is the case in opaque writing systems. If orthographic representations are generated even when the risk of generating an incorrect representation is high, this could mean that generating preliminary orthographic representations during spoken word learning is somehow beneficial for the cognitive system. Indeed, the incidental finding from Experiment 1 showing a positive link between participants' tendency to generate orthographic skeletons and later recall of words with unpreferred spellings, points out to a potentially beneficial role of generating orthographic expediencies in spoken word learning.

2.1.1 Orthographic (in)consistency

The term orthographic (in)consisteny is often considered as being unidirectional, since its most common meaning refers to the degree of print-to-speech (ir)regularity. Nevertheless, languages vary not only in how consistent spelling-to-sound mappings are (feedforward consistency; see Figure I.3) but in how phonology is represented in spelling as well (Schmalz, Marinus, Coltheart, & Castles, 2015). As a result, these two types of (in)consistencies differentiate between languages that are: inconsistent for both spelling and reading (e.g., English; see Ziegler, Stone, & Jacobs, 1997), inconsistent for spelling but consistent for reading (e.g., French; see Ziegler, Jacobs, & Stone, 1996), consistent for spelling but inconsistent for reading (e.g., Danish; Elbro & Pallesen, 2002) and finally, languages that are consistent for both spelling and reading (e.g., Spanish; Defior, Martos, & Cary, 2002).

Figure 2.1: Sound-to-Spelling and Spelling-to-Sound Mappings Across Languages



On a scale comprising both spelling-to-sound as well as sound-to-spelling (in)consistencies, French is situated between Spanish (double consistency) and English (double inconsistency), since it is (in)consistent in only one way.

According to this classification Spanish and English (languages in which the orthographic skeleton hypothesis has been tested so far) vary on two dimensions: spelling-to-sound and sound-to-spelling consistency (see Figure 2.1). By contrast, French is consistent on one, but inconsistent on the other dimension: Compared to Spanish, French orthography is *highly inconsistent* for spelling (sound-to-spelling inconsistency). Since many phonemes map onto more than one grapheme (e.g., the vowel /o/ has at least three possible grapheme representations <o>, <au> and <eau>) predicting the correct spelling of a newly acquired French word comes with higher degree of uncertainty than in Spanish. This property makes French more similar to English when it comes to spelling. At the same time, French is *highly consistent* for reading (spelling-to-sound consistency). Indeed, relying exclusively on grapheme-to-phoneme conversion rules when decoding a novel word in French will in the majority of cases lead to the correct pronunciation. This property makes French more similar to Spanish when it comes to reading. Therefore, while English represents a case of double inconsistency, since it is inconsistent for both spelling and reading, and Spanish represents an example of double consistency, French is only inconsistent in one direction. This one-direction inconsistency makes French a perfect candidate for investigating whether orthographic skeletons are generated even when predicting the correct spelling comes with a higher degree of uncertainty, as it allows us to control for spelling-to-sound consistency.

2.1.2 The present study

Experiment 2 had two goals. As a first, and main goal, we set out to test whether orthographic skeletons are generated even when, due to the properties of the writing system, predicting the correct spelling comes with a greater risk of error. The second, and exploratory goal of the study, was to further inspect the observed link between individual tendency to generate orthographic skeletons during spoken word learning and later recall of these newly acquired words.

2.1.2.1 Main goal of Experiment 2 By testing adults with French as their first and dominant language on exactly the same experimental paradigm used with Spanish speakers, we were able to investigate whether orthographic representations for novel spoken words with multiple spellings are generated even under higher degree of uncertainty. The study had exactly the same structure as the one conducted with L1 speakers of Spanish: A group of French speakers was first asked to spell all novel words that were to be used in the experiment (target and filler novel words). Two weeks later, without knowing that the two experimental sessions were associated, they were trained on a set of 24 novel words. After this aural training, trained and untrained words were presented in short sentences. Importantly, the words participants had been trained on were presented in their unique or one of their two possible spellings (preferred or unpreferred).

Based on both the results from Experiment 1 as well as those reported by Wegener et al. (2018; 2020), we predicted one of the following two outcomes:

1) if orthographic skeletons are generated even for words with more than one possible spelling, reading times for previously acquired words with a unique and those presented in their preferred spellings should not differ. Trained words presented in their unpreferred spellings should however, yield longer reading times. This slowing down should occur since there will be a mismatch between the orthographic skeleton participants generate and the spellings they will later be presented with;

2) if, however, orthographic skeletons are generated only when there is low risk of generating an incorrect expectation (i.e., only in fairly transparent languages), aurally trained consistent French words should be read faster than the inconsistent ones altogether;

Importantly, in both cases, there should be no differences in reading words from the untrained set, thus yielding a significant interaction between training and spelling. This interaction should, in line with Wegener et al. (2018; 2020), be driven by a facilitation present when reading previously acquired words (only consistent or both consistent and preferred). Alternatively, as in Experiment 1, the same interaction could also stem from longer reading times (i.e., inhibitory effect) present only for inconsistent spellings (all or only unpreferred inconsistent).

2.1.2.2 Exploratory goal of Experiment 2 Our secondary goal was to further explore the link between participants' tendency to generate orthographic skeletons and their performance on the final picture naming task. To better understand this link, we looked at whether it is modulated by another construct related to spoken word learning, namely, the phonological short-term memory (PSTM) capacity. Previous research shows that better PSTM capacity, which represents a person's ability to temporarily store phonological information, is associated with better spoken word learning both in children (Gathercole, 2006; Gathercole & Baddeley, 1989) as well as adults (Baddeley, Papagno, & Vallar, 1988). However, the link between PSTM and spoken word learning declines as children get older and gain access to more advanced word learning mechanisms (e.g., comparison with similarly sounding familiar words; see Gathercole, 2006). Interestingly, this decline also overlaps with the official onset of reading. We therefore hypothesized that the decline could, at least partly, be due to the development of the mechanism by which orthographic skeletons are generated. As orthographic knowledge increases, and with it the ability to generate orthographic skeletons, PSTM may be less involved in word learning.

Based on the previous literature, we expected participants with better PSTM capacity to recall better the names of the novel objects at the end of the experiment. At the same time, we expected to replicate the positive correlation between the OSE and word recall. However, if generating orthographic skeletons is indeed modulating the relationship between the PSTM and spoken word learning, a significant interaction between PSTM and

OSE in predicting word recall should be observed. The two in combination should thus be important predictors of word recall. Alternatively, tendency to generate orthographic skeletons and PSTM capacity may be independent from each other. They may actually represent somewhat compensatory mechanisms available once reading has been acquired. If this is indeed the case, there should be no interaction between the two in predicting word recall. Any positive correlation observed between generating orthographic skeletons and word recall should be independent from the positive correlation between PSTM and word recall.

2.2 Methods

2.2.1 Participants

Following the same criteria as in Experiment 1 (see Section 1.2.1), we aimed to collect usable data from approximately 48 participants. However, although a total of 55 participants completed two experimental sessions, due to technical issues with their internet connection, data from five participants failed to be transferred to the server. Additional four participants were removed due to low accuracy in the learning phase (<70%).[6] Therefore, we present data from 46 (39 female) participants aged between 18 and 35 years (M = 22.8, SD = 2.63) who completed both sessions within 10 to 14 days delay between them. Participants were recruited via announcements posted on social media (e.g., student Facebook groups and Twitter). Participants were given a detailed description of the study before confirming their participation. All participants had French as their first and dominant language and all completed a questionnaire on their language skills and habits before starting Session 1 (see Table 2.1 for more details on participants self-reported measures of proficiency). The experiment was approved by the BCBL Ethics Review Board (approval number 060420MK) and complied with the guidelines of the Helsinki Declaration. Participants' written consent was collected at the beginning of each experimental session.

[6] The same cut-off score as in Experiment 1.

Table 2.1: Participant Profile Experiment 2

	Mean	**SD**	**Range**
AoA	0	0	0-0
Self-rated proficiency (0-10)			
Speaking	9.33	0.79	7-10
Understanding	9.61	0.576	8-10
Writing	8.93	0.998	7-10
Reading	9.5	0.753	8-10

Note. Some participants had some basic knowledge of English, but were not proficient in any other language apart from French.

2.2.2 Stimuli

2.2.2.1 Novel Words Two sets of 24 five-phoneme-long CVCVC disyllabic French-like novel words were created (as in Experiment 1, they were classified as set A and B). Each set consisted of eight consistent and 16 inconsistent items (see Table 2.2). Consistent items contained phonemes with one dominant grapheme representation in French (hereafter consistent phonemes) and therefore they all had one dominant spelling (e.g., word /tunav/ as <tounave>). Initial phonemes of all inconsistent words had two possible grapheme representations while the remaining four phonemes were all consistent. Namely, the following four phonemes were used: /ʒ/ which can be written as either <g> or <j>, /k/ which followed by /i/ or /e/ maps onto either <qu> or <k>, /f/ which can be represented with a grapheme <f> or <ph>, and /s/, which followed by /i/ or /e/ can be written with either <c> or <s>. As a result, all inconsistent words had two dominant spellings.

To make sure that consistent items would all be spelled in the same way (e.g., /tunav ə/ spelled as <tounave>), while the inconsistent would be spelled in one of their two dominant spellings (e.g., /ʒebinə/ would be spelled as either <gébine> or <jébine>), ten pilot participants who did not take part in the study, completed a pretest spelling task and confirmed the consistent versus inconsistent spelling manipulation.

Table 2.2: Novel Words from Experiment 2

Set	Consistent	Inconsistent Preferred	Inconsistent Unpreferred
A	/bemanə/	/ʒebinə/	/ʒevabə/
	/danyvə/	/ʒinavə/	/ʒitymə/
	/tunavə/	/sedunə/	/semivə/
	/mabynə/	/simybə/	/sibavə/
	/nypinə/	/fapyvə/	/fanynə/
	/vetagə/	/fedinə/	/fenɔgə/
	/pivadə/	/kemagə/	/kepydə/
	/lybavə/	/kityvə/	/kidunə/
B	/badivə/	/ʒedavə/	/ʒenyvə/
	/devabə/	/ʒimunə/	/ʒitɔgə/
	/mevinə/	/sepidə/	/sebavə/
	/nemunə/	/sitavə/	/sidynə/
	/tabynə/	/fabogə/	/fapunə/
	/pinagə/	/fenybə/	/febadə/
	/lapyvə/	/kenivə/	/kepanə/
	/vinyvə/	/kipynə/	/kimavə/

Note. Words from the inconsistent preferred group were later shown in each participant's preferred spelling whereas words from the inconsistent unpreferred group were presented in the unpreferred spelling.

Consistent words from the two sets were matched on the number of orthographic neighbors (Set A: M = .250, SD = .707 and Set B: M = .375, SD = .744, $t(14)$ = -.344, p = .736), while for the inconsistent items, we made sure that neither of the possible spellings had more than two orthographic neighbors. All words were recorded by a female L1 speaker of French in a sound attenuated cabin using Marantz PMD661.

2.2.2.2 Novel objects The same 48 pictures selected from The Novel Object and Unusual Name (NOUN) Database (Horst & Hout, 2016) and used in Experiment 1 served as novel objects participants were presented with in Experiment 2. Namely, the same pictures paired with Spanish items (see Table 1.2) were matched with their French counterparts (see Table 2.2; i.e., objects paired with consistent items from set A in the Spanish study were paired with consistent items from set A in the French study as well).

Figure 2.2: Novel Objects from Experiment 2

An example object from each set (set A and set B) and spelling group (Consistent, Preferred, and Unpreferred).

2.2.2.3 Sentences from the self-paced reading task Sentences used in the self-paced reading task had the same structure as those used in Experiment 1 (see Table 1.3). They were all four-to-seven words long, and each made a reference to the object preceding it and whose name was to appear within the sentence. Sentences were matched with Spanish sentences on total number of words as well as the position where the target word appeared (see Table 2.3).

Table 2.3: French Sentences used in the Self-Paced Reading Task

Original sentence	**English translation**
Ce xxx est petit	This xxx is small
Ce grand xxx est joli	This big xxx is pretty
Ceci est un xxx gigantesque	This is a large xxx
Ceci est un petit xxx magnifique	This is a small fantastic xxx
Cet objet est un xxx minuscule	This object is a small xxx
Cet objet est un petit xxx magnifique	This object is one small fantastic xxx
Ce grand objet est un xxx magnifique	This big object is one fantastic xxx
Ce grand objet est un magnifique xxx	This large object is a fantastic xxx

Note. Due to syntactic differences across languages the position of the target word is not equivalent in the French sentences and their English translations. Exes represent the place where target words appeared.

2.2.3 Procedure

The study consisted of two experimental sessions completed with a 10-14 days delay between them. In Session 1, participants completed a total of five tasks always presented in a fixed order. The session started with a pre-test spelling task during which participants provided their spelling preferences for all novel words (trained, untrained and fillers). They then completed two distractor linguistic tasks (i.e., a lexical decision and a word spelling task) whose role was to mask the spelling manipulation. Finally, they completed two task aimed at assessing phonological skills: a phoneme deletion task and a pseudoword repetition task. Note that for the present study, only data from the pre-test spelling and the pseudoword repetition are relevant. The other tasks will thus not be discussed further.

Session 2 had exactly the same structure as in Experiment 1. Participants first acquired a set of novel spoken words in the aural training task and were then presented with the words' spellings in the self-paced reading task. The session ended with the picture naming task in which participants had to type the names of all previously acquired objects.

Given that the detailed descriptions of the pre-test spelling task (see

Section 1.2.3.1), the aural training (see Section 1.2.3.2), the self-paced reading task (see Section 1.2.3.3) and the picture naming task (see Section 1.2.3.4) have been provided in Chapter 1, and that the phoneme deletion task was performed as part of another project, only the pseudoword repetition task will be presented in detail here.

2.2.3.1 Pseudoword repetition The aim of the pseudoword repetition task was to obtain a measure of participants' phonological short-term memory capacity and explore whether it correlates with participants' tendency to generate orthographic skeletons. In this task, participants were presented with sequences of monosyllabic French-like nonwords that they first had to repeat in the same order (i.e., starting from the first until the last pseudoword of the sequence; hereafter *forward repetition block*) and then in the reversed order of presentation (i.e., starting from the last and going to the first pseudoword from the sequence; hereafter *backward repetition block*). The order of presentation of the two blocks was fixed (the forward block was always followed by the backward repetition block) and both blocks started with a sequence containing two pseudowords. The sequences gradually increased in length, going from two to eight, and two sequences of each length were presented (i.e., two sequences of two pseudowords, two sequences of three pseudowords, etc.). To make sure that pseudowords within each list sound as phonologically distinct as possible from all the other pseudowords within the same sequence, they all contained a different vowel sound, and none started or ended with the same consonant (the complete list of pseudowords used in the task is presented in Table A.5 in Appendix A.5). Along with the Orthographic Skeleton Effect score (i.e., the OSE score; see Section 1.2.4.2) and the data from the picture naming task (see Section 1.2.4.2), scores from pseudoword repetition task were used to explore the relationship between individual tendency to generate orthographic skeletons and later novel word recall.

Each trial had the following structure. First, a fixation cross appeared at the centre of the screen. After 1000ms a picture of an ear was shown. This picture indicated that the listening part of the trial was in course. At the same time, the sequence of pseudowords was played. In each sequence, there was a 750ms pause between any two pseudowords. After the last pseudoword of the sequence had been played, the picture of the ear disappeared and there was a 500ms pause before the picture of a mouth appeared. This initiated the microphone and the production part of the trial. Participants had 10 seconds to respond before the picture of the mouth disappeared signaling the end of that trial. Participants had to click on the button presented at the centre of the screen to start the next trial and hear the next sequence. At the beginning of each block two practice examples were completed. During the practice trials participants were prompted to record their response and then compare it to the correct one that was played to them after they had given their response. The entire task took participants approximately 20 minutes to complete.

2.2.4 Data pre-processing and analysis

Data obtained in the pre-test spelling task were used to determine the spellings of preferred and unpreferred items in the self-paced reading task. Spelling preferences per item are shown in Appendix A.6 and will no be discussed further. Data from the aural training task served as an indicator of how well participants acquired the novel words they had been trained on and thus exclude those participants with accuracy below the a priori set criteria (<70% of accuracy). Consequently, only descriptive statistics for the aural training are presented. Finally, in the nonword repetition task, a total number of correctly repeated nonwords in both the forward and the backward block was calculated for each participant, and this score was then used as an indicator of their PSTM capacity. Along with the index of participants tendency to generate orthographic skeletons (OSE; see Section 1.3.3), PSTM scores were used to predict word recall in the final picture naming task (i.e., the inverse ALINE score; see section 1.2.4.2).

2.2.4.1 Self-paced reading task RTs were analysed following the procedure described in the Section 1.2.4.1. Namely, before the analysis, raw RTs were visually inspected and data points outside of range (i.e., RTs below 100ms and above 2500ms; eight data points in total) were removed. As indicated by the Box-Cox test (Box & Cox, 1964) the RTs were then log-transformed. Finally, outlier removal was based on scaled residuals. In consequence, only statistical models performed on data without any absolute values of the scaled residuals greater than 3 are reported (1.18% of all data points were removed).

The fixed effects structure was the same as in Experiment 1: The three level factor spelling was deviation coded to create two contrasts of interest: *Spelling1vs2* (consistent spellings coded as -0.33, inconsistent preferred as 0.67 and inconsistent unpreferred as -0.33) and *Spelling1vs3* contrast (consistent spellings coded as -0.33, inconsistent preferred as -0.33 and inconsistent unpreferred as 0.67). The two-level factor *Training* was initially deviation coded (trained words were coded as 0.5 and untrained as -0.5) but then treatment coded in case of a significant interaction (reference level was coded as 0, and the nonreference as 1). Finally, the covariate *Set* was also deviation coded (Set A as -0.5 and Set B as 0.5).

Again, in order to avoid overfitting the models which would lead to reduced statistical power (Matuschek et al., 2017), random-effects structure was build following a parsimonious data-driven approach (Bates et al., 2015). As a result, all reported models represent the highest nonsingular converging models and include the maximal random effects structure for all experimental manipulations of interest supported by the data (see Table 2.5 for exact model structure).

2.3 Results

2.3.1 Aural training

The accuracy in all four blocks of the aural training, as well as the final check phase, was high (see Table 2.4). Importantly, there were no differences between the two sets of words in the final check phase (Set A: $M = 93$, $SD = 7.08$; Set B: $M = 93.8$, $SD = 6.37$; $t(44) = .394$, $p = .695$).

Table 2.4: Accuracy in the French Aural Training Task

	Block1	**Block2**	**Block3**	**Block4**	**Final Check**
Set A	95.8 (4.64)	96.9 (2.49)	97.1 (4.29)	95.6 (4.61)	93 (7.08)
Set B	96.4 (4.55)	98.1 (3.19)	95.5 (7.35)	95.5 (5.69)	93.8 (6.37)

Note. Mean percentage of accuracy (SDs) per training block and in the final block of the aural trainig task.

2.3.2 Self-paced reading

In the self-paced reading task, participants read both trained and untrained words presented in their unique (i.e., consistent words), preferred or unpreferred spellings (i.e., inconsistent words). Mean RTs for all target words, measured relative to the onset of the word at the center of the screen, are shown Figure 2.3 and the distributions of RTs are shown in Appendix A.7 (see Figure A7).

The full model structure including both fixed and random effects is shown in Table 2.5. The model looking into both trained and untrained words showed no effects of either the *Spelling1vs2* ($\beta = .002$, $SE = .027$, $t = .082$, $p = .935$) or the *Spelling1vs3* contrast ($\beta = .034$, $SE = .028$, $t = 1.22$, $p = .227$). The main effect of *Training* was however significant ($\beta = -.044$, $SE = .019$, $t = -2.24$, $p = .031$), indicating that the trained words overall ($M = 518$, $SD = 255$) were read faster than the untrained ones ($M = 554$, $SD = 291$). Importantly, while the interaction between *Training* and *Spelling1vs2* was not significant ($\beta = .034$, $SE = .037$, $t = .917$, $p = .364$), the interaction between *Training* and *Spelling1vs3* was ($\beta = .077$, $SE = .037$, $t = 2.05$, $p = .046$).

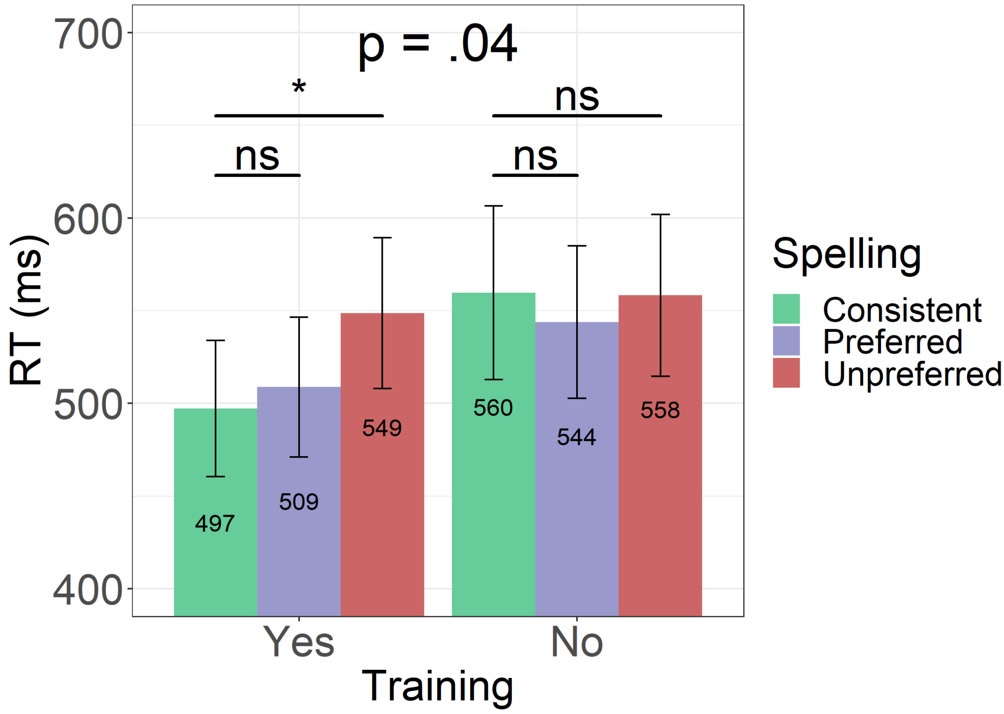
Table 2.5: Full Structure of the Model from Experiment 2

Fixed effects	**Estimate**	**SE**	***t* value**	***p***
(Intercept)	6.12	0.068	90.3	**< .001*****
Training	-0.044	0.019	-2.24	**< .05***
Spelling1vs2	0.002	0.027	0.082	0.935
Spelling1vs3	0.034	0.028	1.22	0.227
Set	0.211	0.135	1.56	0.125
Training: Spelling1vs2	0.034	0.037	0.917	0.364
Training: Spelling1vs3	0.077	0.037	2.05	**< .05***
Random effects	**Variance**		**Std. Dev.**	
Participant (Intercept)	0.206		0.453	
Participant: Training (slope)	0.007		0.086	
Participant: Spelling1vs3 (slope)	0.004		0.060	
Item (Intercept)	0.003		0.059	
Item: Training (slop)	0.002		0.044	

Note. The full model looks into both trained and untrained words.

Model looking into trained words only (by treatment coding the factor training; trained coded as 0 and untrained as 1), showed that the difference between consistent and unpreferred spellings was significant (β = .073, *SE* = .032, *t* = 2.25, *p* = .027). The same difference however, was not significant when only untrained words were considered (β = -.003, *SE* = .034, *t* = -.111, *p* = .912).

Figure 2.3: Reaction Times From the Experiment 2



Reaction times for both trained (Yes) and untrained (No) consistent, preferred and unpreferred word spellings. Error bars represent the standard error of the mean.

To summarise, the analysis of RTs showed a significant effect of aural training given that trained words overall yielded faster RTs as compared to the untrained ones. Importantly, this effect was driven by faster RTs for consistent and inconsistent preferred spellings.

2.3.2.1 Further inspection of the training effect To further inspect the significant training effect in both the present, as well as the study conducted with Spanish speakers, two models (one per language) looking into the effect of training for each spelling condition separately were run. A custom contrast coding was used to create three contrasts of interest: the first one compared trained and untrained consistent spellings (hereafter *ConsistentSpelling* contrast; consistent trained coded as -0.5, consistent untrained coded as 0.5, the rest coded as 0), the second one compared trained and untrained preferred spellings (hereafter *PreferredSpelling* contrast; preferred trained coded as -0.5, preferred untrained coded as 0.5, the rest coded as 0), and the third one compared trained and untrained unpreferred spellings (hereafter *UnpreferredSpelling* contrast; unpreferred trained coded as -0.5, unpreferred untrained coded as 0.5, the rest coded as 0).

The model looking into data from French speakers included by-participants and by-item intercepts as well as by-participant random slopes for *PreferredSpelling* contrast. The model showed a significant effect of training for consistent spellings (β = .081, SE = .024, t = 3.35, p = .001), but not for preferred (β = .047, SE = .025, t = 1.89, p = .065) or unpreferred ones (β = .004, SE = .024, t = .175, p = .861).

The same model run with data obtained from Spanish speakers included by-participants and by-item intercepts as well as by-participant random slopes for *ConsistentSpelling* and *UnpreferredSpelling* contrasts. The model showed no differences between trained and untrained consistent spellings (β = -.019, *SE* = .027, *t* = -.694, *p* = .491) nor a difference between trained and untrained preferred spellings (β = -.036, *SE* = .025, *t* = -1.46, *p* = .144). However, the difference between reading trained and untrained unpreferred spellings was significant (β = -.063, *SE* = .030, *t* = -2.09, *p* = .042).

2.3.3 Exploratory analysis: The role of phonological memory in generating orthographic skeletons

To investigate the relationship between phonological short-term memory (PSTM) measured through pseudoword repetition task, and participants' tendency to generate orthographic skeletons (OSE) on novel word retention, a multiple regression with Inverse ALINE score as a dependent variable was performed. The analysis included two predictors as well as an interaction between them: *PSTM score*, measured as the total number of correctly repeated pseudowords in both the forward and the backward pseudoword repetition task (ranging from 34 to 122 of correct responses) and the *OSE score*, operationalized as the difference between reading aurally acquired inconsistent unpreferred and consistent words (ranging from -217ms to

475ms). Both predictors were standardized before running the model. The collinearity between them was checked with VIF.mer (Frank, 2011) and all VIFs were below 2. Finally, the model with the best fit (i.e., the model with the highest adjusted R-squared value) was the one that included both predictors along with their interaction.

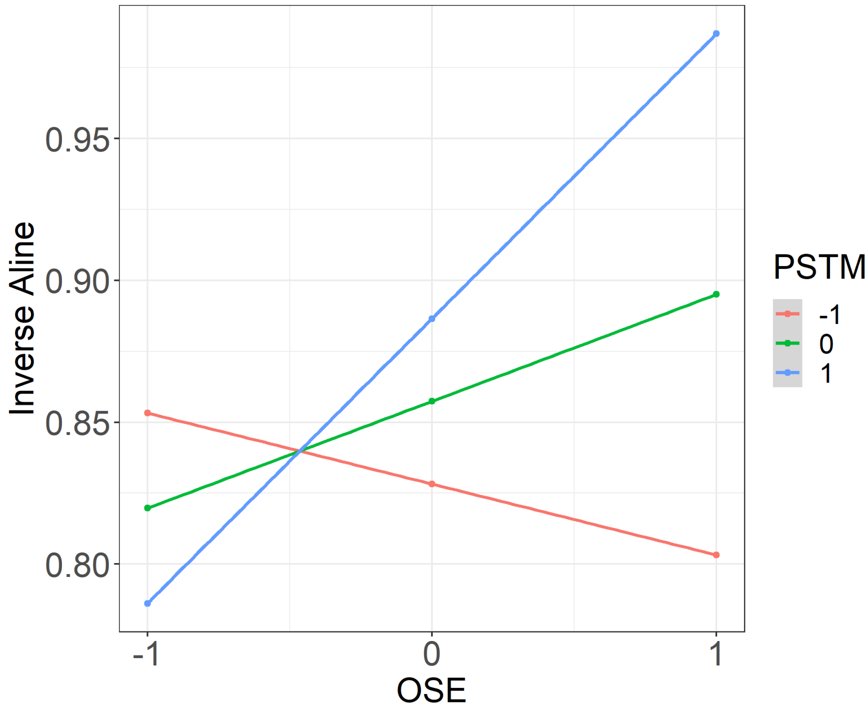
The results of the multiple regression show that the two predictors and their interaction predict 28.8% of the variance in the picture naming task (R^2= .28, $F(3,39) = 5.12$, $p < .01$; see Table 2.6). OSE on its own significantly predicted the inverse ALINE score (β = .038, SE = .018, t = 2.07, p = .045), indicating that higher values of OSE lead to better recall (i.e., a higher score in the picture naming task). By contrast, the PSTM score failed to reach the level of significance (β = .029, SE = .015, t = 1.96, p = .057). However, there was a significant interaction between the two (β = .063, SE = .024, t = 2.58, p = .014). The interaction suggests that the relationship between one of the predictors and the outcome, is moderated by the other predictor. For instance, the link between OSE and the inverse ALINE score, changes as a function of the PSTM score: for low values of PSTM the increase in OSE does not lead to differences in the outcome variable (i.e., the inverse ALINE), while for the higher values in PSTM the values of the outcome increase as the OSE scores increase (see the partial plot presented in Figure 2.4 for the visual representation of the significant interaction between the predictors).

Table 2.6: Regression Model Predicting the Performance on the PNT

	Inverse ALINE score			
Predictors	*Estimates*	*SE*	*Statistic*	*p*
Intercept	0.86	0.01	61.14	**<0.001**
OSE	0.04	0.02	2.07	**0.045**
PSTM	0.03	0.01	1.96	0.058
OSE: PSTM	0.06	0.02	2.58	**0.014**

R2 / R2 adjusted 0.282 / 0.227

Figure 2.4: Interaction between PSTM and OSE



The moderating effect of the OSE on the relation between phonological short-term memory and word retention (Inverse ALINE similarity score). Both predictors (PSTM and OSE) are represented as z-scores.

2.4 Discussion

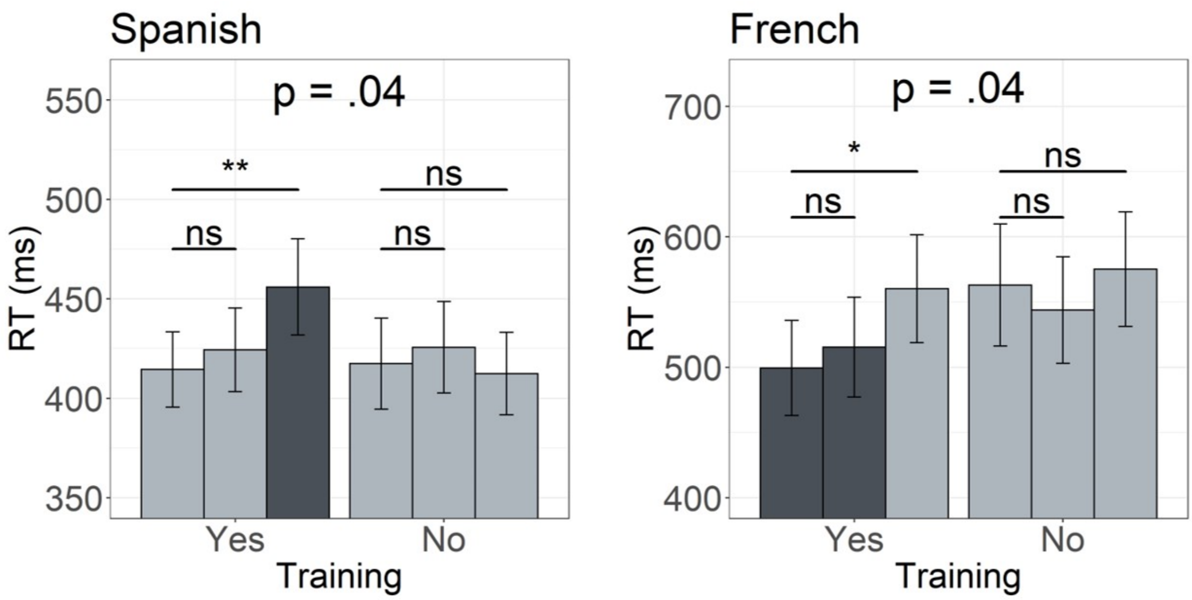
Experiment 2 investigated whether orthographic skeletons for words with multiple spellings are generated even in a language in which the overall probability of generating an incorrect representation is high. To test whether results from Experiment 1 can be generalised to a language with highly inconsistent sound-to-spelling mappings, we looked at how adult French speakers read novel words previously acquired through aural training. As in Experiment 1, novel words were shown in their unique or one of the two possible spellings. Overall, the results of Experiment 2 show that French speakers generated orthographic skeletons for newly acquired spoken words. This conclusion was supported by the following observations: Firstly, trained words altogether were read faster than the untrained ones. As in Wegener et al. (2018) study, this aural training advantage was driven by faster reading times for words shown in spellings matching participants' expectations. Secondly, words presented in the unpreferred of the two possible spellings led to significantly longer reading times as compared to words with a unique spelling. This indicated that there was a mismatch between the orthographic skeletons participants had generated as a result of aural training, and the spellings they were presented with in the subsequent reading task. Finally, no significant differences in reading times were observed for untrained words shown in their unique, preferred and unpreferred spellings. These results thus show that orthographic skeletons for words with more than one spelling are indeed generated even in a language with high degree of sound-to-spelling inconsistency.

As previously discussed, in languages like French, the relationship between sounds and letters is complex since many sounds have more than one grapheme representation (Ziegler, Jacobs, & Stone, 1996). Due to the higher degree of uncertainty, and consequently higher risk of generating incorrect representations, the process of generating orthographic skeletons for newly acquired spoken words may be more conservative (i.e., restricted to situations when only one spelling is possible), or may even be suppressed in speakers of opaque languages. Here however, we show that this is not the case. French speakers exhibited a similar pattern of results as previously tested Spanish speakers. In both studies, inconsistent words presented in their unpreferred spellings yielded significantly longer reading times as compared to words with only one possible spelling. This demonstrates that French speakers were not affected by the overall complexity of sound-to-spelling mappings present in the French writing system. Furthermore, findings from French speakers are in line with those reported by Wegener and colleagues (Wegener et al., 2018; Wegener et al., 2020; see also Beyersmann et al., 2020). Across two studies, conducted with English-speaking children, Wegener and colleagues observed a significant reading advantage for previously trained novel words as compared to the untrained words. Crucially, this facilitation was driven by faster reading times found only for predictable but not for previously acquired words with unpredictable spellings. Given that the latter had more than one possible spelling, it remained unclear whether orthographic skeletons in opaque languages are generated even when, due to multiple options, the spelling of the novel spoken word is uncertain. Using the same experimental design as the one employed with Spanish speakers allowed us to control for the number of inconsistent spellings and demonstrate that orthographic skeletons for words with multiple, and hence highly unpredictable spellings, are indeed generated in languages with overall high sound-to-spelling uncertainty. The findings from the present study thus suggest that generating orthographic skeletons may be beneficial when learning novel spoken words, since it is done even when the risk of error is high.

2.4.1 Comparison with results obtained from Spanish speakers

Even though the findings from the present study are in line with those observed with Spanish speakers, an unexpected difference between the two experiments is the reversed effect of training, and in particular, the significant aural *training advantage* observed with French, but not with Spanish speakers. The significant training effect found in the present study stems from overall faster reading times observed for trained as compared to untrained words. This advantage in reading aurally familiar words is in line with previous research showing facilitatory effects of aural training on subsequent word reading, as well as the formation of novel orthographic representations (e.g., Álvarez-Cañizo, Suárez-Coalla, & Cuetos, 2019; Johnston et al., 2004; McKague, Pratt, & Johnston, 2001, Wegener et al., 2018, Wegener et al., 2020). The evidence for the orthographic skeleton hypothesis is thus found in the *facilitatory effect* present when reading previously acquired words shown in correctly predicted and expected spellings (see Figure 2.5). In the study conducted with Spanish speakers however, expected spellings did not differ in reading times across the two training conditions. The significant training effect was actually driven by longer reading times present only for unpreferred spellings. This slowing down thus resulted in a *training disadvantage* for words that did not match participants' spelling expectations. Based on this, we can conclude that evidence for the orthographic skeleton hypothesis in Spanish is manifested as an *inhibitory effect* (as opposed to facilitatory) (see Figure 2.5). In other words, while Spanish speakers were slower to read previously acquired words with unexpected spellings, French participants were faster to read previously acquired words with expected spellings (see dark grey bars on Figure 2.5).

Figure 2.5: Comparison between Spanish and French Results



These differences probably stem from different experiences speakers of opaque and transparent languages have with written language. While the former are used to encountering words with multiple possible spellings, and may therefore be *positively* surprised whenever their expectations are confirmed, the latter (e.g., Spanish speakers) rarely have the occasion to see multiple spellings for one phonological word form. As a result, they may be more *negatively* surprised when encountering spellings not in line with their expectations. Moreover, this cross-linguistic comparison, showing facilitatory or inhibitory effects of generating orthographic skeletons, suggests that generating orthographic skeletons may come with a *reading benefit* in opaque (facilitatory effect), but not in transparent languages (inhibitory effect).

2.4.2 The link between generating orthographic skeletons on word learning

In addition to showing that generating orthographic skeletons has important consequences for subsequent word reading – either by facilitating the reading of expected spellings or by slowing down the reading of the unexpected ones – the present study reveals some consequences for spoken word learning linked to the process of generating orthographic skeletons. As seen in Experiment 1 (see Section 1.3.3), generating orthographic skeletons is associated with better recall of words shown in their unpreferred spellings: participants with larger orthographic skeleton effects were the ones who recalled more correctly the novel words presented in their unpreferred spellings. To further examine, and this way better understand the observed link between individual tendency to generate orthographic skeletons and later word recall, we set out to relate this finding to the previous literature. Based on the previous research showing that the positive link between PSTM skills and spoken word learning (Baddeley, Papagno, & Vallar, 1988; Gathercole, 2006; Gathercole & Baddeley, 1989; Gathercole, Willis, Emslie, & Baddeley, 1992) declines around the age of 8 (Gathercole, Willis, Emslie, & Baddeley, 1992), we wanted to see whether generating orthographic skeletons is at least partly related to this decline. As a result, we included a measure of PSTM in the experiment and then ran a multiple regression model predicting the outcome in the picture naming task.

The link between PSTM and the performance on the picture naming task was not significant, this way failing to replicate the results reported in the literature. However, we once again observed a positive correlation between the OSE score, representing participants' tendency to generate orthographic skeletons, and naming accuracy. Participants who were more strongly prone to generate orthographic skeletons, recalled better the words at the end of the experiment. Importantly, there was a significant interaction between the OSE score and PSTM capacity in predicting the accuracy in the picture naming task. This interaction indicated that the link between participants' tendency to generate orthographic skeletons and word learning was indeed modulated by PSTM: only participants with higher PSTM scores were able to benefit from generating orthographic skeletons. In other words, participants who obtained higher accuracy scores at the end of the experiment (i.e., those who remembered more correctly the words they had been trained on) did not only have higher PSTM capacity score but higher OSE scores as well. This interaction thus suggests that both generating orthographic skeletons as well as having good PSTM capacity are important for recalling (and therefore learning) novel spoken words. In line with reading time data, this finding shows that generating orthographic skeletons comes with a benefit, a *word learning* benefit in this case. Although it may be limited to skilled readers only, this findings shows that orthography is involved in spoken word learning even when it is not present nor relevant for the learning process.

2.4.3 Conclusion

In summary, data presented so far show that orthographic skeletons are generated even for words with multiple spellings, and that, both in languages with high (Spanish), as well as those with low sound-to-spelling consistency (French). However, we demonstrated that the consequences of generating orthographic skeletons for subsequent reading are manifested differently depending on the overall consistency of the language. Although both studies expand our understanding of the mechanism driving sound-to-spelling conversions, the nature of the mechanism by which phonological representations are converted into orthographic ones has yet to be described. Specifically, it remains unclear whether orthographic skeletons are generated unconsciously, as an automatic response to acquiring novel phonological word forms (M. Johnston, Castles, et al., 2003; Tyler & Burnham, 2006) or whether generating them represents a strategy participants employ with the aim to facilitate word learning. All previous studies testing the orthographic skeleton account employed word learning paradigms in which participants were explicitly instructed to learn novel words. Since orthography has been shown to have a facilitatory role in word learning (Rosenthal & Ehri, 2008), generating orthographic skeletons could indeed have served as a mnemonic tool helping participants in the acquisition of novel words. To adjudicate between the two possibilities and this way test whether orthographic skeletons are generated automatically during novel word learning, or whether generating them represents a strategy participants consciously employ in order to support word memorization, the process of generating orthographic skeletons should be compared in explicit and implicit learning contexts. This was indeed done in the final study of the thesis (see Experiment 3).

3 The nature of orthographic effects in spoken word learning

Work presented in this chapter is based on:

Jevtović, M., Kapnoula, E.C., and Martin, C.D. (in preparation). Implicit versus explicit creation of spelling expectations during auditory word learning

3.1 Introduction

So far we have presented data showing that aural training with novel spoken words leads to the creation of orthographic skeletons (i.e., preliminary orthographic representations). In addition, we demonstrated that, although they have different consequences for subsequent word reading, orthographic skeletons are generated both in a language with relatively low level of spelling uncertainty (e.g., Spanish; see Experiment 1) as well as in a language with higher level of spelling uncertainty (e.g., French; see Experiment 2). In Spanish, a language with highly consistent sound-to-spelling mappings, generating orthographic skeletons leads to an *inhibitory effect*. Conversely, generating such representations in a language with overall higher probability of generating an incorrect representation (due to inconsistent sound-to-spelling mappings), seems to be beneficial as it results in a *facilitatory effect*. These findings showing that orthographic skeletons are generated even when there is a risk of error, could be interpreted as showing that engaging in the process of generating orthographic skeletons does not entail a high cognitive cost, and may therefore be automatic in nature. However, these two experiments were not designed to test the nature of the mechanism by which orthographic skeletons are generated. Therefore, it remains unclear whether this mechanism is automatic or maybe even voluntary in nature. In the final experimental chapter of this thesis we directly test the nature of the mechanism responsible for sound-to-spelling conversion during auditory word learning. We do so by testing whether generating orthographic skeletons results from an unconscious, and hence automatic process, or whether orthographic skeletons are generated as a result of a strategic process consciously initiated by participants with the aim to facilitate novel word learning. Understating the nature of this converting mechanism would not only add to the orthographic skeleton account, but would help us better understand the nature of the orthographic influence on spoken language processing.

As discussed in the Introduction, studies looking into the role of orthography in word learning, mainly focused on exploring whether presenting orthographic labels along with novel phonological forms has a facilitatory effect on the learning process. In these studies, novel words' spellings were thus explicitly presented during the learning process, and participants could use them as an additional memory cue easing the acquisition of novel vocabulary. Ehri and Wilce (1979) were the first to show that explicitly linking orthographic labels to novel phonological representations comes with a learning benefit in early readers (see also Rosenthal & Ehri, 2008). In their study, first and second grade children were both faster and more accurate to acquire novel vocabulary when both spelling and sound were present during the learning phase. These findings were later expanded to both older children (i.e., eight and nine-year-olds; see Ricketts, Bishop & Nation, 2009) as well as adults (Nelson et al., 2005). Interestingly, the *orthographic facilitation* observed by Ehri and Wilce (1979) occurred not only when spellings were explicitly shown during the learning phase, but when children were instructed to imagine novel words' spellings as well. To explain this learning advantage present in the absence of orthography, the authors argued that the instructions children had received prompted them to generate words' spellings by forming visual images during the learning phase. In return, these visual images, in the form of words' spellings, facilitated the acquisition of novel phonological forms just as if they had been explicitly present during the acquisition process.

With these findings in mind, and in particular, the one showing that orthographic facilitation occurs even in the absence of orthography, one could argue that the effects observed both in our, as well as the two studies conducted by Wegener et al. (Wegener et al., 2018, 2020), resulted from a strategic process participants had initiated with the aim to facilitate the learning process. Although orthography was not present during the learning process in the previous two experiments, and participants were not told they would later be presented with words' spellings, they were supposed to learn the novel phonological word forms they were presented with. That is, participants in both of these studies knew they were in a situation of *explicit learning*. Consequently, they could have relied on their knowledge of phoneme-to-grapheme correspondences to generate additional learning cues, in the form of a preliminary orthographic representations. These orthographic representations would then help them learn the novel words. By contrast, based on the reasons described in the previous paragraph, and which have to do with cross-linguistic differences observed so far, the other possibility according to which orthographic skeletons result from an automatic and involuntary process is also plausible. This is indeed what some studies showing orthographic effects during spoken language processing would argue for. Specifically, since in some tasks relying on one's orthographic knowledge actually comes with a disadvantage (Castles et al., 2003), orthographic effects may very well be involuntary (i.e., automatic). To adjudicate between these two possibilities, we conducted a study in which the creation of orthographic skeletons was compared in an active versus passive word learning context (i.e., explicit versus implicit learning). Implicit learning defined as the type of learning that occurs without conscious awareness (Reber & Winter, 1994) does not require attention to learn. Seeing evidence for the orthographic skeleton hypothesis in a passive learning context would mean that generating orthographic skeletons is an automatic rather than strategic process.

3.1.1 The present study

The goal of Experiment 3 was to test whether generating orthographic skeletons during spoken word learning occurs automatically (i.e., without participants' conscious attention) or whether generating them results from a voluntary process participants purposely engage in with the aim to facilitate the learning process. To that end, two groups of Spanish speakers were tested reading previously acquired spoken words with either one, or two possible spellings (consistent and inconsistent words). However, while half of the participants learned the novel words through explicit learning (hereafter *active learners*), the other half was naive regarding the goal of the study and thus acquired the words implicitly (hereafter *passive learners*). Importantly, both groups completed exactly the same auditory word learning task, the only difference being the instructions they received: passive learners were told that the task was testing their ability to recognize objects presented in different sizes and coulours, while the active learners were instructed to learn all the words as their performance would later be tested (see Section 3.2.3.2).

As in the previous experiments, all participants first provided their preferred spellings for all novel words to be used in the main task (those they would later be trained on, as well as those from the untrained set). Two weeks later, they completed the aural training task right before being exposed to novel words' spellings in a self-paced reading task. Upon completing the reading task, they went through a learning check. Note that in contrast to the first two experiments, which assessed participants' learning accuracy before the reading task, the learning check was completed after the reading task. This was done so as not to reveal the aim of the study to the group of passive learners (i.e., if they were to see the learning check before reading, they could have realized what the aim of the study was).

Based on the findings from the experiment conducted with Spanish speakers, in the group of active learners we expected to observe a significant interaction between spelling and training driven by longer reading times only for previously acquired words shown in their unpreferred spellings. This would imply that participants had generated orthographic skeletons for all previously acquired words. That is, the pattern of results in the group of active learners should replicate the one observed in Experiment 1. Importantly however, if the process underlying the creation of orthographic skeletons is automatic (i.e., not driven by participants' strategies) the same pattern of results should be present in the group of passive learners. By contrast, if orthographic skeletons are a result of a conscious and intentional process, the group of passive learners should not show any differences in reading previously acquired consistent and inconsistent words with unique, preferred or unpreferred spellings.

Note that we were not interested in any interactions between the two groups. Testing two groups was however needed in order to make sure that the paradigm used to test the creation of orthographic skeletons in a passive learning context is indeed appropriate and can detect any potential effects of previous aural training. Therefore, all our planned comparisons concern per group analysis.

3.2 Methods

3.2.1 Participants

A total of 64 participants with Spanish as their first and dominant language were tested in the study (see Table 3.1 for complete information about participants linguistic profile). They were randomly split in two groups (33 active and 31 passive learners). The groups were matched on their working memory capacity (M_{active} = 1.94, SD_{active} = 2.74; $M_{passive}$ = 2.32, $SD_{passive}$ = 1.94; $t(62)$ = -.642, p = .523) as well as nonverbal IQ (M_{active} = 85.8, SD_{active} = 7.95; $M_{passive}$ = 85.5, $SD_{passive}$ = 10.7; $t(62)$ = .089, p = .929).

Participants were recruited from the BCBL Participa database and received 20 euros for their participation in the study. The experiment was approved by the BCBL Ethics Review Board (approval number 011221SM) and complied with the guidelines of the Helsinki Declaration. All participants gave their written consent at the beginning of each of the two sessions.

Table 3.1: Participant Profile Experiment 3

	Mean	SD	Range
AoA	0	0	0-0
Picture naming (0-65)	64.8	0.46	63-65
LexTale (0-100%)	94.1	5.25	73.33-100
Interview (1-5)	5	0	5-5
Self-rated proficiency (0-10)			
Speaking	9.78	0.45	8-10
Understanding	9.68	0.562	8-10
Writing	9.66	0.59	7-10
Reading	9.73	0.51	8-10

Note. As in Experiment 1, some participants had some knowledge of a second or even a third language but none of them was highly proficient in any language other than Spanish.
[a] There are a total of 65 pictures to be named in the BEST (making 65 the maximum possible score). [b] Self-rated proficiency data are missing for three participants.

3.2.2 Stimuli

The same novel words from Experiment 1, along with the pictures associated with them, were used in the study (see Section 1.2.2). However, in order to minimize misspellings (e.g., participants writing <chuñe> instead of <yuñe> or <lluñe> for the item /ʎuɲe/), as well as empty responses (i.e., participants leaving out the target phoneme <uñe>) observed in the pre-test spelling task from Experiment 1 (see Appendix A.2), a novel set of audio recordings with improved quality was employed. As in Experiment 1, the recordings were made by an L1 Spanish speaker coming from the same region as the participants tested in the experiment.

3.2.3 Procedure

The experiment was organised in two sessions. During Session 1, participants first provided their preferred spelling options for all novel words to appear in the experiment in a pre-test spelling task. They then completed two linguistic distractor tasks (i.e., a lexical decision and a real word spelling task). Next, to make sure the two groups (i.e., the active and passive learners) were matched on their working memory capacity, all participants completed a working memory N-back task, followed by a pseudoword repetition task. To ensure participants were also matched on their non-verbal IQ, at the end of Session 1 they completed a three minute Raven's Progressive Matrices task (Raven et al., 2003).

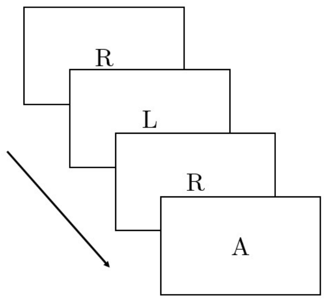
Session 2 started with the aural training task which was the same for both groups, the only exception were thus the instructions given to the participants. Right after the aural training participants completed a short distractor task (i.e., the Simon task; Simon, 1969). Next, they were presented with novel words' spellings in a self-paced sentence reading task. Upon completing the reading task, participants were asked several questions regarding their impressions of the study. Importantly, to check whether passive learners were really naive as to the goal of the experiment, all participants were explicitly asked to describe what they think the study was investigating. To make the experiment as similar as possible to the previous two, following this short questionnaire, participants completed a picture naming task and a learning check.

An important improvement in regards to the first two experiments was that at the beginning of each session, participants completed a short headphone-check task (Woods, Siegel, Traer, & McDermott, 2017). The aim of this task was to make sure participants wear headphones throughout the experiment. On each trial they heard three consecutive sounds, and had to indicate which one was the quietest. Differentiating the loudness of the three sounds is relatively easy with headphones, but almost impossible to do without them. There were in total 6 trials and participants had to correctly respond to at least four in order to start the experiment. Otherwise, they would repeat the headphone check until reaching this criteria.

Since most of the tasks have already been presented in the previous chapters (see Section 1.2.3.1 for a description of the pre-test spelling task, and Section 1.2.3.3 for the self-paced sentence reading task) and given that nonword repetition task will not be discussed further, here we present only the tasks specific to this experiment.

3.2.3.1 N-back task To make sure that participants from the two groups were matched on their working memory capacity, they all completed a 2-back letter task. In this task participants were presented with a sequence of letters presented one at a time, and their task was to determine whether the letter on the screen matches the letter that appeared two trials (letters) before (see Figure 3.1). Participants were instructed to respond by pressing the button 'M' to the relevant letter and to withhold their responses to distractor letters. Each letter stayed on the screen for 500ms and participants had additional 2000ms (blank screen) to give their response before the appearance of the next letter. There were in total 26 main and 10 practice trials. The main dependent variable was the number of correctly identified 2-back matching items.

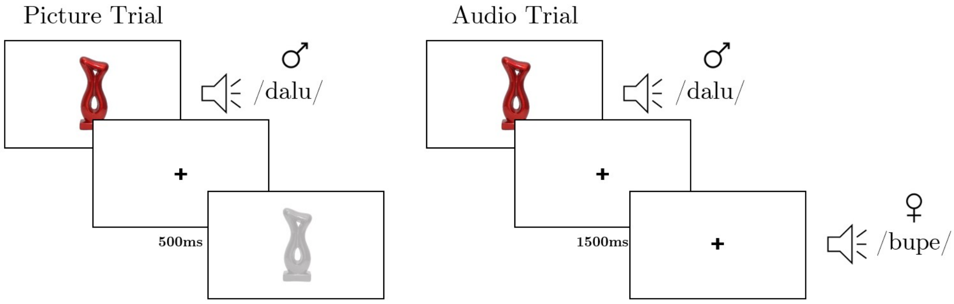
Figure 3.1: N-back task



Participants had to press 'M' on their keyboard only when the letter presented matched the letter presented two trials ago. In this example participants had to press 'M' upon the appearance of the second R.

3.2.3.2 Aural training Both groups completed the same aural training task in which pictures of 24 novel objects were presented. The task had the structure of a match-mismatch task, and each trial consisted of two parts: a *presentation part* on which a picture of a novel object was presented at the center of the screen while its name was played (in a male voice), and a *verification part* on which either a matching or a mismatching picture or word (but not both), were presented. Participants' task was to respond whether the item matched the previously presented picture or name. To make the task more engaging for the participants, as well as force them to process both the pictures and the auditory words, pictures and words presented on the verification part of the trials varied in physical properties from the pictures and words presented on the presentation part (see Figure 3.2). Namely, pictures were always presented in grey scale, while auditory words were played in a different gender voice (i.e., in a female voice). Another manipulation consisted of presenting smaller and larger versions of the pictures (50% smaller or larger than the original size), while the same manipulation for objects' names consisted of presenting the female voice either in higher (70dB) or lower intensity (60dB) as compared to the male voice during the presentation part. These changes in physical properties of both pictures and auditory words were equally distributed over all trials. Finally, to indirectly force participants to keep the picture-word associations in their memory, the interstimulus interval between the presentation and the verification part of each trial was jittered from 500ms to 1500ms in 500ms steps. Apart from making the task more engaging, not knowing when the verification part would start (due to the unpredictable interstimulus interval) was supposed to prompt the participants to keep the picture-word association in their working memory until giving their response.

Figure 3.2: Trials of the Aural Training Blocks



The structure of the phonological training task. An example of a picture-word pair followed by a matching picture (on the left) and a picture-word pair followed by a mismatching word (on the right).

As in the previous experiments, to limit the learning load, novel objects were presented in four blocks of six. All blocks contained 75% of *match trials* (same picture or word as on the presentation part of the trial) and 25% of *mismatch trials* (different picture or word from those on the presentation part of the trial). This proportion was chosen so not to make the task, as well as learning picture-word associations, too difficult. In particular, it could be argued that presenting different pictures and words during one trial (as is the case on mismatch trials) is hindering the learning process. Therefore, within the blocks, each picture-word pair was followed nine times by the same picture or word - *match trials* - and three times by a different picture or word - *mismatch trials*, giving this way a total of 72 trials per block. After completing all four blocks, participants completed a final fifth block with the same structure (i.e., picture-word pairs were followed by either matching or mismatching pictures or words), in which all 24 pictures were presented. However, there was no manipulation in the sound intensity nor the picture size. Each of the 24 pictures was presented on four trials (two match and two mismatch trials) giving this way a total of 96 trials. The duration of the entire task was around 45 minutes.

3.2.4 Learning check

To have a measure of how well participants learned the picture-word associations, at the end of the experiment they completed a short learning check. On each trial four pictures were presented on the screen (two on the left and the right side of the screen and two pictures on the upper and lower part of the screen; see Figure 1.3) and the name of only one of them was played. Participants' task was to choose the object which corresponded to the name that had been played. Each object was paired only once with its name, giving this way a total of 24 trials.

3.2.5 Data pre-processing and analysis

RT data from the two groups were analysed separately, since a priori we were not interested in any group interactions. Therefore, we first present data from active learners and then those from passive learners. In both tasks, RTs from the self-paced reading task were analysed following the procedure described in Section 1.2.4.1. Consequently, prior to any statistical analysis, RTs were inspected and extreme values removed (i.e., RTs below 100ms and above 2500ms; 11 data points in active and 10 data point in passive learners). In line with the Box-Cox test, RTs were then log-transformed.

Moreover, given that despite the fact that novel words had been re-recorded to make the sound clearer, participants still made errors for items starting with the sound /ʎ/. That is, some participants provided neither of the two target grapheme representations (i.e., <y> or <ll>) but a representation which does not correspond to the target sound (e.g., <hi>). Those trials for both trained and untrained items were therefore removed before the analysis (3.03% of all data points in the group of active learners and 2.55% of all data points in the group of passive learners).

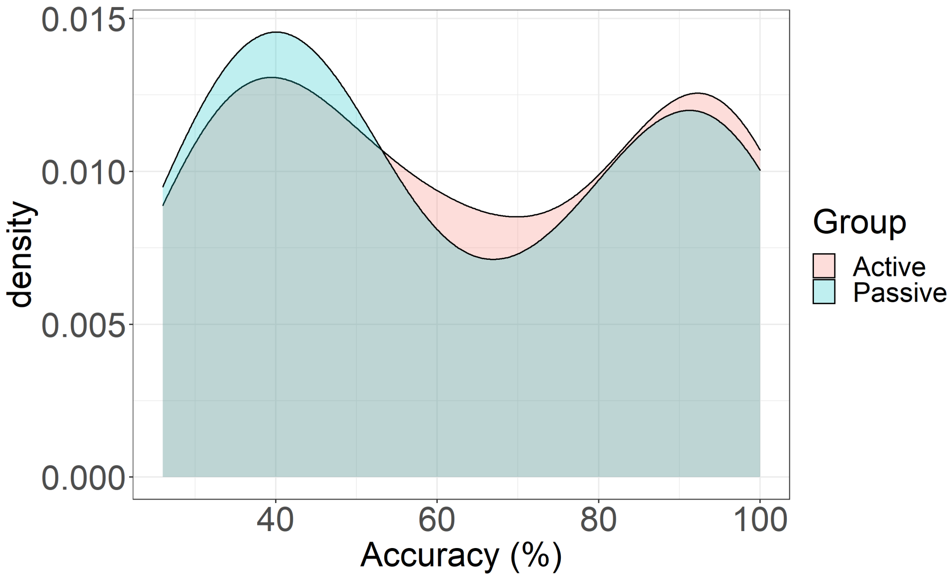
3.3 Results

Preferred spellings from the pre-test spelling task for both word sets (Set A and B) are shown in Tables A.8 and A.9 in Appendix A.8.

3.3.1 Learning check

The overall accuracy from the learning check phase did not differ between the groups ($M_{active} = 64.1$, $SD_{active} = 26$; $M_{passive} = 62.6$, $SD_{passive} = 26.4$; $t(62) = .226$, $p = .822$), suggesting that the type of instruction (explicit versus implicit) did not lead to different learning outcomes. Distribution of accuracy for both active and passive learners is shown in Figure 3.3.

Figure 3.3: Distribution of Accuracy in the Learning Check



Distribution of accuracy for the two groups of participants (active learners are shown in pink and passive learners in blue. The two groups did not differ in their performance on the learning check task.

3.3.2 Self-paced reading

3.3.2.1 Active learners Mean RTs for the three different spellings of both trained and untrained words, observed in the group of active learners, are shown in Figure 3.4. Distributions of RTs are shown in Appendix A.9 (see Figure A9.1).

The model with the maximal random-effects structure justified by the data (Bates et al., 2015), looking at both trained and untrained words, included by-participant random intercepts as well as by-participant random slopes for the factor training and the *Spelling1vs3* contrast (see Table 3.2).

The model showed no effect of *Training* (β = -.023, *SE* = .020, *t* = -1.18, *p* = .243), or *Spelling1vs2* contrast (β = .029, *SE* = .019, *t* = 1.55, *p* = .120). However the *Spelling1vs3* contrast, indicating a difference in reading unpreferred and unique spellings, was significant (β = .045, *SE* = .020, *t* = 2.21, *p* = .033). Contrary to Experiment 1, there was no interaction between *Training* and *Spelling1vs3* contrast (β = -.038, *SE* = .039, *t* = -.987, *p* = .323). The same was the case for the interaction between *Training* and *Spelling1vs2* contrast (β = -.038, *SE* = .038, *t* = -1.00, *p* = .316).

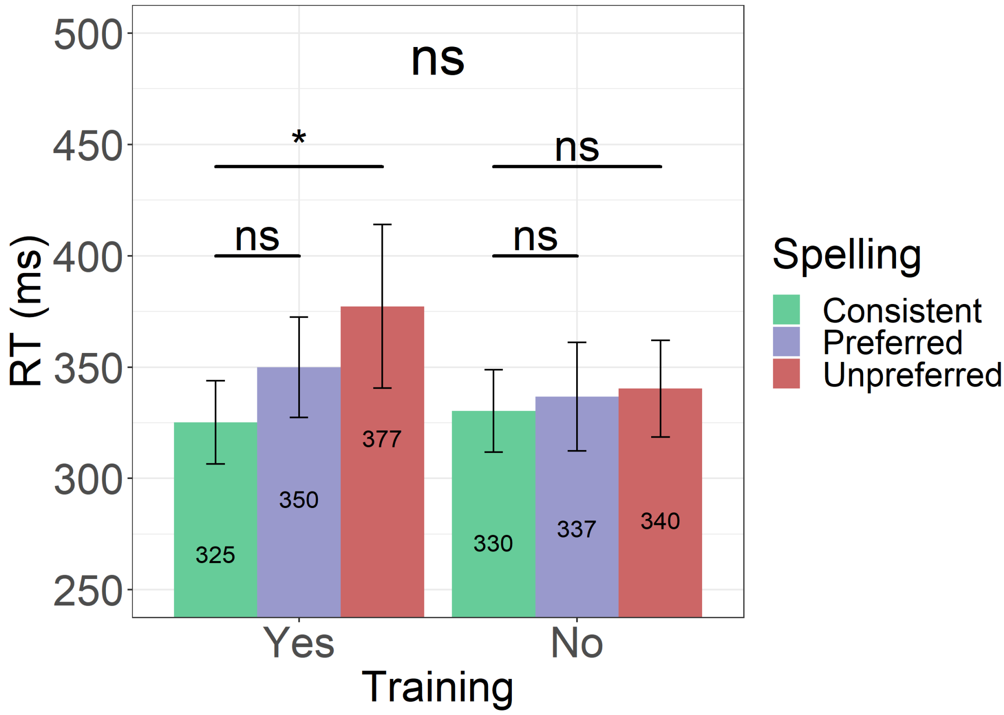
Table 3.2: Full Structure of the Model looking into Active Learning Group

Fixed effects	**Estimate**	**SE**	**t value**	**p**
(Intercept)	5.71	0.058	98	< .001***
Training	-0.023	0.020	-1.19	0.243
Spelling1vs2	0.029	0.019	1.55	0.120
Spelling1vs3	0.045	0.020	2.21	< 0.05*
Set	-0.149	0.116	-1.28	0.210
Training: Spelling1vs2	-0.038	0.038	-1.00	0.316
Training: Spelling1vs3	-0.038	0.039	-0.987	0.323
Random effects	**Variance**		**Std. Dev.**	
Participant (Intercept)	0.105		0.324	
Participant: Training (slope)	0.004		0.068	
Participant: Spelling1vs3 (slope)	0.001		0.033	

Note. The full model looks into both trained and untrained words.

Although the interaction between training and *Spelling1vs3* contrast was not significant, we ran two additional models: one looking only into trained (trained words coded as 0) and the other looking only into untrained words (untrained words coded as 0), and found that the *Spelling1vs3* contrast was significant within the trained (β = .064, *SE* = .028, *t* = 2.28, *p* = .024), but not within the untrained words (β = .025, *SE* = .028, *t* = .916, *p* = .361).

Figure 3.4: Reaction Times in the Active Learning Group From Experiment 3



Reaction times for both trained (Yes) and untrained (No) consistent, preferred and unpreferred word spellings. Error bars represent the standard error of the mean, and the numbers within the bars represent the mean RTs for each condition.

3.3.2.2 Passive learners Mean RTs for the three different spellings of both trained and untrained words, observed in the group of passive learners, are shown Figure 3.5 and the distributions of RTs are shown in Appendix A.9 (see Figure A9.2).

The main model looking into both trained and untrained words (see Table 3.3) showed no main effect of *Training* (β = -.011, *SE* = .021, *t* = -.505, *p* = .617), *Spelling1vs2* contrast (β = -.002, *SE* = .022, *t* = -.113, *p* = .910) or *Spelling1vs3* contrast (β = .021, *SE* = .035, *t* = .831, *p* = .406). Neither of the two interactions: *Training* and *Spelling1vs2* contrast (β = .003, *SE* = .043, *t* = .069, *p* = .945) or *Training* and *Spelling1vs2* contrast were significant (β = .016, *SE* = .043, *t* = .366, *p* = .714).

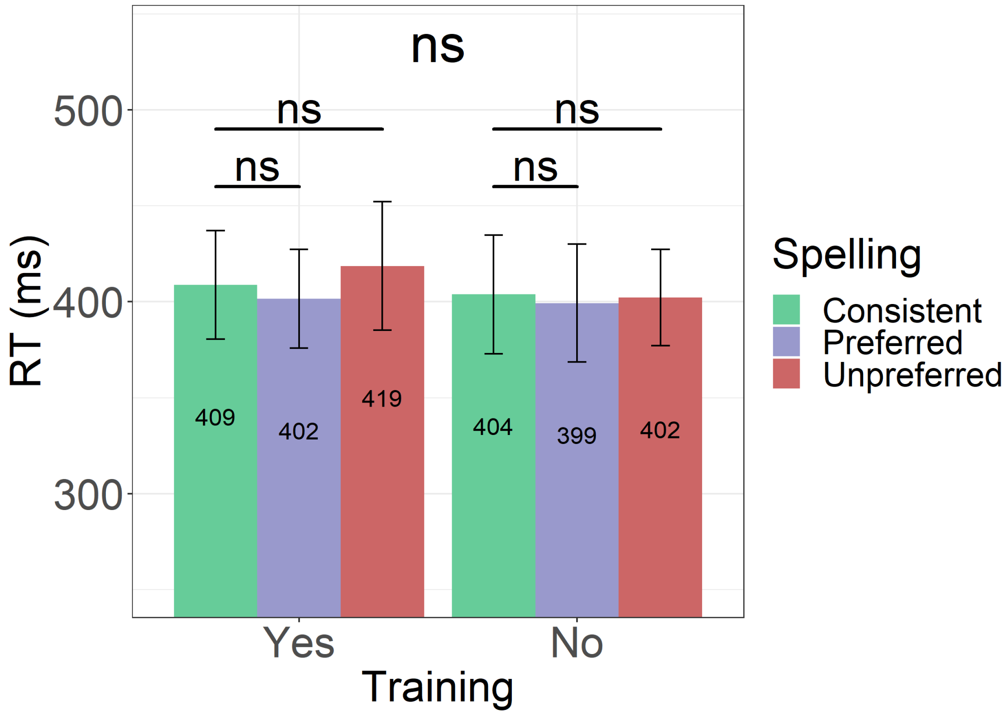
Table 3.3: Full Structure of the Model looking into Passive Learning Group

Fixed effects	Estimate	SE	t value	p
(Intercept)	5.89	0.072	98	$< .001^{***}$
Training	-0.013	0.022	-0.599	0.554
Spelling1vs2	0.004	0.021	0.182	0.856
Spelling1vs3	0.021	0.021	0.997	0.319
Set	0.059	0.144	0.412	0.684
Training: Spelling1vs2	0.004	0.042	0.089	0.928
Training: Spelling1vs3	0.010	0.042	0.245	0.807
Random effects	**Variance**		**Std. Dev.**	
Participant (Intercept)	0.121		0.348	
Participant: Training (slope)	0.005		0.072	

Note. The full model looks into both trained and untrained words.

Moreover, no differences in RTs were found within trained words: *Spelling1vs2* (β = .002, *SE* = .030, t = .065, p = .948), *Spelling1vs3* (β = .016, *SE* = .030, t = .532, p = .595), or the untrained ones: *Spelling1vs2* (β = .006, *SE* = .030, t = .191, p = .848) and *Spelling1vs3* (β = .026, *SE* = .030, t = .876, p = .381).

Figure 3.5: Reaction Times in the Passive Learning Group from the Experiment 3



Reaction times for for both trained (Yes) and untrained (No) consistent, preferred and unpreferred word spellings. Error bars represent the standard error of the mean, and the numbers within the bars represent the means for each condition.

3.3.2.3 Exploratory analysis To better understand the pattern of results from the two learning groups, three additional exploratory analysis (not planned apriori) were conducted. The first one compared RTs for unique, preferred and unpreferred spellings across the two groups of learners within the same model. The second one compared the data from Experiment 1 to those of active learners from the present study. In addition, to explore whether the absence of a significant interaction between *Training* and *Spelling1vs3* contrast in the group of active learners is due to the lack of statistical power in the present study, we ran a power analysis based on the estimates from Experiment 1. This enabled us to see how many participants would be needed to detect a significant interaction in a future study.[7]

[7]Note that we are calculating power for a *future study* based on the estimates from Experiment 1, and not the power of the Experiment 1. The latter is pointless given that the *p* values are already known (Hoenig & Heisey, 2001).

3.3.2.3.1 Combined group analysis To test whether there is an effect of group, and precisely, type of learning, on reading previously acquired spoken words with unique, preferred and unpreferred spellings, a combined analysis with both groups of learners was performed. The fixed factor *Learning Group* was deviation coded (active learners coded as -0.5, passive learners coded as 0.5).[8]

Apart from the effect of *Learning Group* (β = -.003, SE = .0009, t = -2.71, p = .007), showing that reading times in the group of active learners (M = 348, SD = 234) were overall faster than those observed in the group of passive learners (M = 404, SD = 258), none of the main effects or interactions were found to be significant ($p > .05$).

3.3.2.3.2 Comparison with Experiment 1 To compare reading latencies for trained and untrained unique, preferred and unpreferred words observed in the group of active learners from the present study to those from Experiment 1, the same model as the one described above, but with the factor *Experiment* (Experiment 1 coded as -0.5, Experiment 2 coded as 0.5) was run.

The model showed a significant effect of *Experiment* (β = -.228, SE = .075, t = -3.032, p = .003) since overall reading times for active learners were faster in the present (M = 348, SD = 234) as compared to the first study (M = 421, SD = 209). In addition, the model detected the effect of *Spelling1vs3* contrast (β = .041, SE = .015, t = 2.78, p = .006). None of the interactions were significant ($p > .05$).

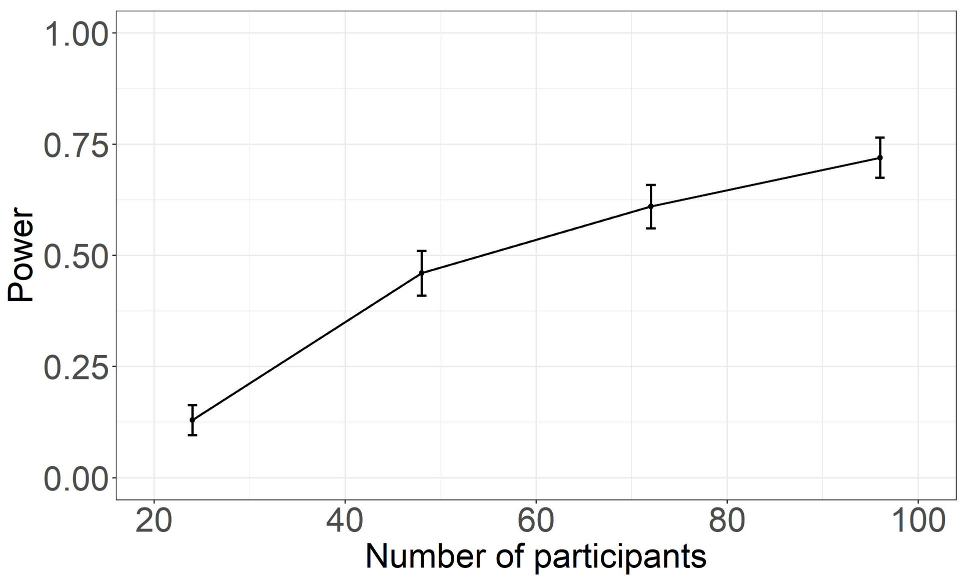
3.3.2.3.3 Power analysis based on data from Experiment 1 The power analysis based on the data from Experiment 1 was performed by creating a dataframe with the same design as in Experiment 1, using the *designr* package (Rabe et al., 2021) in R. Next, 400 novel data sets based on the estimates obtained in Experiment 1 (100 per sample size, ranging

[8]The model had the following structure: logRT~1+Training:Spelling1vs2:LearningGroup + Training:Spelling1vs3:LearningGroup + Set + (1+training_coded+group3_1||participant)

from 24 to 96 in steps of 24) were simulated, and the model with the maximal random-effects structure (including both by-participants and by-item random intercepts, by-participants random slopes for training, two contrasts of interest and their interaction with training, as well as by-item random slops for the factor training) was run on each of the datasets. The power per sample size was then calculated as the percentage of times a significant interaction between *Spelling1vs3* contrast and *Training* was obtained (e.g., obtaining a significant interaction 60 out of 100 times would indicate a power of 0.6 for that particular sample size).

Power for a potential future study as a function of sample size (from 24 to 96 in steps of 24) is shown in Figure 3.6.

Figure 3.6: Power as a Function of the Sample Size



Power analysis was performed on the data from Experiment 1.

Based on this power analysis, we conclude that in order to detect the significant interaction between *Spelling1vs3* contrast and *Training* with a power of 0.7 we would need at least 96 participants, which is almost three times more participants than tested in the present study.

3.4 Discussion

The final experiment of the present thesis further investigated the mechanism by which novel phonological representations are converted into orthographic ones during auditory word learning. In particular, we tested whether the process of generating orthographic skeletons is unconscious and automatic, implying that it functions any time a novel word is acquired, or whether it is voluntary in nature. The latter would suggest that participants engage in generating orthographic skeletons consciously every time they need to learn a new spoken word. To that end, two groups of Spanish speakers were tested reading previously acquired spoken words with a unique or two possible spellings. Importantly, the groups differed in whether they acquired novel words through explicit or implicit instruction. That is, although the aural training task was exactly the same for both groups, *active learners* - just like participants tested in Experiments 1 and 2 - received explicit instructions and were aware that they had to learn novel picture-word associations. *Passive learners* by contrast, completed the same aural training task thinking their perception of novel objects was being tested. These differences in participants' awareness of the learning process allowed us to test the automaticity of the mechanism responsible for the creation of orthographic skeletons.

Overall, the results show differences in reading aurally trained words between the two groups of learners. While no significant effects in reading aurally trained words with unique, preferred or unpreferred spellings were found in the passive learning group, active learners showed a similar pattern to the one reported in Experiment 1. Even though no interaction between spelling and training was found (as was the case in both Experiment 1 and 2), the difference between reading words with unique as compared to those presented in the unpreferred of the two spellings was significant only within the set of trained words. We take this difference within trained words as evidence that active learners did generate orthographic skeletons, and this, regardless of the number of possible spellings. In line with the orthographic facilitation literature showing positive impact of orthography novel word learning, active learners may have consciously chosen to do generate orthographic skeletons in order to facilitate the word learning process. By contrast, given that participants from the passive learning group did not show differences in reading previously acquired words with a unique or two possible spellings, we do not have reasons to believe that they had generated orthographic skeletons as a result of the aural training. Importantly, since they seem to have learned the novel words equally good as active learners, their newly acquired word representations were probably entirely phonological in nature. Having no explicit reason to generate orthographic representations in addition to the phonological ones, passive learners did not automatically activate their orthographic knowledge during the training phase.

Apart from providing more insight into the nature of the process responsible for generating orthographic skeletons, findings from the present experiment expand on the previous research reporting orthographic effects in spoken language processing. Throughout this thesis we relied on orthographic (in)consistency to test the involvement of orthography in spoken word learning. Along with the studies conducted by Wegener and colleagues (Wegener et al., 2018, 2020) the work presented here thus reveals another way in which orthographic knowledge affects speech processing. By making use of their sound-to-letter mapping knowledge, skilled readers are able to generate preliminary orthographic representations even in the absence of orthography. However, previous literature showing the impact of orthography on speech perception, argues that the orthographic effects are automatic in nature. There is indeed evidence showing that upon hearing spoken words, their orthographic counterpart are automatically co-activated (Chéreau et al., 2007; Perre & Ziegler, 2008). Data from the present experiment by contrast, suggest that in spoken word learning orthographic effects may not be automatic but could actually be driven by a strategic and hence voluntary process. Being confronted with a novel spoken word will not automatically lead to the creation of its orthographic analog. As they are generated by a voluntary process (i.e., via explicit instruction) these preliminary orthographic representations may also not affect auditory perception in the same way already familiar words do. Future research could set out to explore the qualitative difference between orthographic representations generated through visual exposure (i.e., reading) to those generated in the absence of orthography.

To explore the absence of the interaction between training and spelling, we conducted a power analysis on data and estimates from Experiment 1. This way we were able to obtain the information about the statistical power for a future study employing the same design, as a function of the sample size. As shown in Figure 3.6, to detect a significant interaction with a power of 0.7, we would need at least 96 participants, which compared to 32 participants tested in the active learners group, implies that the present study did not have enough power to detect the interaction we were interested in. Another potential reason for the absence of the significant interaction may be linked to differences in learning tasks between this and Experiment 1. Although we tried to make the two studies as similar as possible, we had to come up with a task that would at the same time allow us to test implicit word learning. The task we employed to test both explicit and implicit learning was consequently more difficult than the one used in Experiments 1 and 2. Firstly, the task used in the present study was twice as long as the one used in the first two studies. Moreover, each trial of the aural training task used in the present study consisted of two parts (i.e., presentation and verification part). Finally, the learning task used in the present study was also more engaging, as participants had to monitor and process more aspects of the pictures and words presented to them (e.g., words played in different gender voices, pictures shown in different sizes and colours, etc.). That learning tasks differed in their difficulty is supported by differences in overall accuracy from the learning check task, which was lower here than in both Experiment 1 and 2 (more than 90% in the first two as compared to 62% in the present study). It is important to note however, that while learning check consisted of 144 trials in the first two experiments (each picture was paired with its name on 6 different trials), in the present study each picture was paired with its name only once (24 trials in total). This means that participants in the first two experiments had more occasions to improve their overall accuracy score. As a result, the two tasks are not comparable and we cannot conclude much based on the observed differences in learning scores. Having said that, although the learning accuracy may be considered lower in the present as compared to the previous two studies, it is reassuring that active and passive learners did not differ in their overall learning rate. Given that the participants from the passive learning group did not know they were supposed to learn the picture-word associations, one could expect to see lower learning rates than in the active learning group of participants, since the latter were explicitly instructed to focus on picture-word associations as their performance would later be tested. The fact that no differences were found, and that learning rate was almost the same across the two group (62% compared to 64% in the active learners group) suggests that the paradigm we developed for the purpose of the present study could be used in future research to study mechanisms involved in explicit and implicit word learning.

Although implicit and explicit learning mechanisms *per se* were not of the main interest for the present study, differences and similarities between the two groups of learners should nevertheless be discussed in relation to the previous literature. Based on both behavioural (Sobczak & Gaskell, 2019) as well as neuroimaging data (Batterink & Neville, 2011) showing that acquiring novel words takes more time in implicit than in explicit learning contexts, it was somewhat surprising that after a relatively short aural training, no differences in learning scores were found between the groups. In fact, distributions of the learning scores were almost identical in the two groups. This finding could reflect the failure of explicit learners to properly acquire novel word forms. Alternatively, the absence of a difference between the groups could also stem from methodological issues raised in the previous paragraph. It could be the case that with more learning check trials, a difference between the groups would emerge. Importantly, the findings from the present study add to the previous literature by showing another difference between explicit and implicit learning. While explicit spoken word learning may lead to the creation of (preliminary) orthographic representations, implicit acquisition of novel phonological word form does not involve the activation of orthography.

A somewhat surprising finding was a significant difference in total reading times observed when data from the group of active learners from the present study were compared to those obtained in Experiment 1. Namely, even though the pattern of reading times was similar across the two studies, indicating that participants in the present study, similarly to those tested in Experiment 1, generated orthographic skeletons during the learning phase, active learners from the present experiment were overall faster to read both trained and untrained words. As previously discussed, the two studies employed learning tasks that differed in their overall difficulty and duration. Since the task was more engaging and it took them more time to complete, participants in the present study could have been eager to complete the experiment thus leading to their faster performance on the reading task. However, it is important to keep in mind that the present study included a sample size almost twice as small as the one from Experiment 1. Since the variability in reading times was large (see Figure 3.4), the pattern of results observed in the present study may therefore be influenced by a few extremely fast participants. Testing more participants with the aim to achieve more power and reach at least the same sample size as in Experiment 1, will help clarify reading time differences.

To sum up, the final experiment of the present thesis investigated the nature of the mechanism responsible for the creation of orthographic skeletons during spoken word learning. Although more participants need to be tested in order to draw any strong conclusions about the automatic or voluntary nature of skeleton generating mechanism, data from the present experiment suggest that generating orthographic skeletons may actually occur only in active learning contexts. This means that the effects described so far may not be inherent to the cognitive system but may actually be under voluntary control. Alternatively, the process of generating orthographic skeletons as a result of implicit learning may need more time.

Discussion

The present thesis investigated the role of orthography in novel spoken word learning. Specifically, we asked whether orthographic knowledge, which comprises sound-to-spelling conversion rules, impacts the acquisition of novel phonological representations by converting them into orthographic ones. Stuart and Coltheart (1988) proposed that children can make use of their orthographic knowledge to form preliminary orthographic representations (i.e., orthographic skeletons) for aurally familiar words, and that, even before seeing the words' real spellings. This idea, and in particular, the link between orthographic knowledge and spoken word learning, had been investigated in several studies employing novel word learning paradigms. Johnston and colleagues (2004) for instance, showed that adult English speakers were able to encode initial orthographic representations of newly acquired spoken words. In a masked priming paradigm, which taps into unconscious and automatic processes (Forster & Davis, 1984), they observed significantly faster reading times for aurally acquired words preceded by identical primes (e.g., e.g., <spathe> before <SPATHE>) as compared to those preceded by purely phonological primes (e.g., <spaith> before <SPATHE>). More recently, Wegener and colleagues (2018, 2020) manipulated novel words' spellings, and showed that when first encountered in print, aurally trained words with predictable spellings were read faster than words with unpredictable spellings. By demonstrating that orthographic representations can be generated even in the absence of orthographic input, these studies provided some of the first evidence for the account proposed by Stuart and Coltheart (1988). Nevertheless, they left open several important questions regarding the process underlying the creation of these preliminary orthographic representations. Firstly, from the existing research it remained unclear whether orthographic skeletons are generated for all aurally acquired words, regardless of the number of possible spellings, or only for words with a unique and thus completely predictable spelling. Moreover, as all studies had been done in English, a language with highly inconsistent both spelling-to-sound and sound-to-spelling mappings, one could argue that the observed results may be specific to languages resembling English. Given that English speakers are used to encountering inconsistent and unpredictable spellings, generating orthographic skeletons may be a process specifically developed by their cognitive system in order to ease the reading of new printed words. Finally, and in line with the previous point, none of the previous studies was able to say anything about the nature of the process by which orthographic skeletons are generated. In particular, it remained unclear whether the observed effects were due to participants' personal strategies, or whether they were driven by a process inherent to the cognitive system which gets developed once sound-to-spelling conversion rules had been acquired. These questions were thus explicitly tested in the present thesis.

Across three experiments we showed that aural training with novel word forms prompted skilled readers to use their knowledge of sound-to-spelling mappings to generate preliminary orthographic representations. Importantly, we saw that these orthographic skeletons are generated before the first visual encounter with words' actual spellings. Moreover, we described the conditions under which orthographic skeletons are generated, as well as factors that influence their creation. In the first experiment we showed that orthographic skeletons are generated even when there is uncertainty regarding the correct spelling. When first confronted with novel words' spellings, Spanish skilled readers showed no differences in reading aurally familiar words with unique and those with two possible spellings,. However, this was the case only when the latter were shown in the preferred spelling option. We took this finding as evidence that these Spanish speakers had indeed generated orthographic skeletons for all previously acquired spoken words (both for words with a unique as well as those with two possible spellings).

In the second experiment we replicated the findings from Experiment 1 and showed that orthographic skeletons for words with more than one possible spelling are generated even in a language in which the overall probability of generating an incorrect representation is high (i.e., French). Since the probability of generating an incorrect representation is higher in French than in Spanish, one could argue that generating orthographic representations even under uncertainty (i.e., for words with multiple legal spelling options) may be specific to consistent languages. By replicating the results observed with Spanish speakers, we showed that this is not the case and that orthographic skeletons for words with multiple spellings can indeed be generated even in a highly inconsistent language.

Finally, in Experiment 3 we investigated the nature of the orthographic effects in spoken word learning. More precisely, we asked whether generating orthographic skeletons represents an automatic (unconscious) or voluntary (conscious) process. Although they are not conclusive, due to relatively small sample size, our data suggest that orthographic effects in spoken word learning may very much be driven by participants' intentions, and are therefore voluntary rather than automatic in nature. This finding thus implies that although orthographic skeletons are generated for all aurally acquired words, they may actually be generated only when doing so can help participants in the (explicit) word learning process.

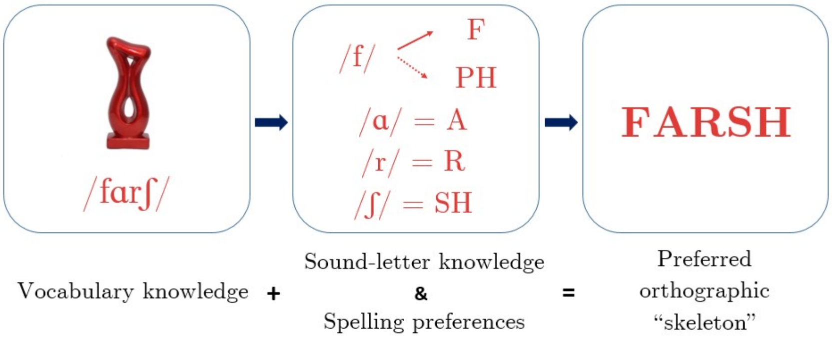
Theoretical and practical implications

Findings from the three experiments conducted in the course of the present thesis have important theoretical and practical implications for both reading development theories, as well as word learning research. We will start by discussing our findings in the light of previous research. Importantly, we will highlight what the present thesis adds to what we already know about the relationship between orthography and phonological processing. Finally, we will propose several ideas on how these findings could be used to bridge the gap between theoretical and practical work in both reading development and novel vocabulary acquisition.

Consequences of generating orthographic skeletons for word reading

The orthographic skeleton hypothesis was initially developed with the aim to explain the observed (positive) link between oral vocabulary knowledge and word reading success (see Wegener, Beyersmann, Wang, & Castles, 2022). According to the previous accounts of reading acquisition and development, orthography exerts its effects on word reading *at the moment* of the first visual encounter with the aurally familiar word's spelling. This occurs either via orthographic learning (Castles & Nation, 2008) or phonological decoding (Share, 1995). The orthographic skeleton hypothesis however, goes a step further by claiming that orthographic influence can begin even *before* the first visual encounter with aurally familiar words' spellings. Through sound-to-letter conversions, aurally familiar words get converted into preliminary orthographic representations. These representations are generated during the learning phase and are then held in lexical memory, similarly to existing orthographic representations of already familiar words. Importantly, these representations are malleable and can be modified upon visual encounter with words' actual spellings (Wegener et al., 2020). Due to the existence of these preliminary orthographic representations, the first time a reader is confronted with a word's spelling, reading should be facilitated, but only when the representation in the lexical memory matches the presented spelling. The work presented here further shows that representations in the lexical memory are influenced by personal spellings preferences rather than statistical properties fo the writing system (see Figure D.1).

Figure D.1: Updated Representation of the Orthographic Skeleton Hypothesis



Having a word in the oral vocabulary and knowing how sounds map onto letters will lead to the creation of a preliminary orthographic representation (i.e., the orthographic skeleton). Importantly, orthographic skeletons for words with more than one possible spelling will be in line with personal spelling preferences.

The present thesis expands on the orthographic skeleton account by showing that the consequences of generating orthographic skeletons for subsequent word reading may actually be dependent on the properties of the writing system, and in particular, the overall sound-to-spelling consistency. The pattern of results observed in Experiment 2, which was conducted with French speakers, follows that of English speakers reported by Wegener et al. (Wegener et al., 2018, 2020). In French and English, which are both opaque languages, reading was facilitated when presented spellings were in line with orthographic skeletons participants had previously generated. The pattern of results observed with speakers of a transparent language such as Spanish (Experiments 1 and 3), was however, somewhat surprising. Namely, the two studies done with Spanish speakers demonstrated that aural training does not necessarily lead to reading facilitation, since no differences in reading trained and untrained words with unique and preferred spellings were observed. This absence of the training advantage in Spanish was explained as stemming, at least partly, from the transparency of the Spanish writing system. As previously discussed, speakers of a transparent language such as Spanish have more experience with encountering spellings in accordance with their expectations. Being confronted with spellings that do not match their predictions could thus lead to slower reading times (i.e., a surprisal effect). By contrast, speakers of opaque languages such as English and French are less often presented with correctly predicted spellings. The surprisal effect could therefore be expressed as a significant facilitation for words that do match their predictions. Based on this reasoning and the work presented here, we can conclude that generating orthographic skeletons is actually more beneficial in inconsistent languages.

The cross-linguistic differences reported in this thesis should thus be incorporated into the existing theoretical accounts aiming to understand the positive, and according to some, causal link between oral vocabulary knowledge and word reading McKague et al. (2001). Work by Wegener et al. (Wegener et al., 2018, 2020), as well as the results from Experiment 2, show that this link could indeed be supported by a mechanism leading to the creation of orthographic skeletons, in speakers of opaque languages. However, data obtained in the two studies done with Spanish adult speakers not only failed to replicate the positive influence of aural training on word reading, but they showed that in a transparent language generating orthographic representations can actually hinder subsequent reading. Although the absence of the aural training advantage may have been caused by experimental material (e.g., relatively simple and short words), we should not exclude the possibility that in adult speakers of transparent languages, aural training does not necessarily come with an advantage. On a more practical level, these findings could be used to improve reading instructions and interventions. Given that generating orthographic skeletons seems to be under voluntary control, both children as well as adults learning to read in an opaque writing system should be encouraged to generate orthographic skeletons not only when learning novel spoken words, but for already familiar words as well. Of course, more research on speakers of languages with consistent language systems (e.g., Italian or German), is needed to confirm these conclusions regarding the observed cross-linguistic differences, and any reading instructions and/or interventions should be designed accordingly.

Consequences of generating orthographic skeletons for word learning

Although the main motivation behind the orthographic skeleton account was not to describe the word learning process per se, but rather its relationship with reading, the present data provide new insights into how novel vocabulary is acquired in the auditory domain. In the first place, along with the two studies conducted by Wegener and colleagues (Wegener et al., 2018, 2020), findings from the present thesis show that once orthographic code is acquired, both children and adults can make use of it during spoken word learning. This means that in addition to learning a novel phonological representation, the concept it refers to, the link between the two and finally integrating the word in the network of already existing words (McMurray et al., 2016), learning a new word may include an additional step. The additional step would consist of generating orthographic analogs of the acquired phonological word forms, but only when the words are acquired via explicit instruction. Our results thereby expand on those reported by Ehri and Wilce (1979). In their seminal study, Ehri and Wilce (1979) showed that linking orthography to novel phonological word forms facilitates the learning process. This orthographic facilitation was present even when participants were not exposed to words' spellings, but were instructed to imagine them. Ehri and Wilce (1979) thus concluded that the instruction prompting participants to imagine words' spellings (i.e., visual representations of novel words) resulted in an additional memory cue aiding the learning process. The work presented here shows that during spoken word learning, participants generate visual representations of novel words - in the form of preliminary orthographic representations - even when they are not explicitly instructed to do so, but when generating spelling expectations can be of help and facilitate the learning process.

By showing differences in reading novel words acquired through explicit as compared to those acquired through implicit aural instruction, findings from the present thesis provide novel insights into the mechanisms underlying these two types of learning. Whereas explicit training with novel spoken words resulted in the creation of orthographic skeletons, the same training in the implicit learning context did not. Previous research reported several advantages of explicit over implicit instruction. Of particular interest for the present thesis are the findings showing faster establishment of novel lexical representations acquired trough explicit learning (Batterink & Neville, 2011; Sobczak & Gaskell, 2019). Since orthographic skeletons represent an additional memory cue promoting the learning process, generating them may be one of the reasons leading to the superiority of the explicit as compared to implicit learning. Future research could thus set out to test the causal role of generating orthographic skeletons in the faster acquisition of novel lexical representations during explicit learning.

Interestingly, our data show that participants vary considerably in their ability to generate orthographic skeletons, since differences in reading unpreferred and unique spellings were not uniform in any of the three experiments. Although the sample size in each of the studies is not large enough to make any strong conclusions regarding the individual differences, the two exploratory analyses we performed with the aim to better understand the consequences of generating orthographic skeletons, show that participants with larger surprisal effects also show better retention of the words they were trained on. In addition, we show that good phonological short-term memory skills alone, which have been so far considered the main predictor of spoken word learning success (Gathercole, 2006), may not be enough to explain individual variability in novel word recall. These individual differences could stem from the quality of orthographic representations (in this case orthographic skeletons) participants generated during the learning task (Perfetti & Hart, 2001; Perfetti, Wlotko, & Hart, 2005). Participants with incomplete or vague orthographic representations had fewer cues to hold on to when recalling the novel words, which would explain their lower accuracy on the picture naming task (see Ehri & Wilce, 1979). This reasoning is supported by a post-hoc exploratory analysis performed on data from Experiment 1. To explore whether participants who were more consistent in their spelling preferences, meaning that they invariably preferred one grapheme over the other (e.g., always chose letter <b> over <v> for inconsistent words starting with the phoneme /b/, or grapheme <ll> over <y> for items whose initial phoneme was /ʎ/) exhibit a somewhat different pattern of results than those participants varying in their preferences (e.g., for some items they chose <b> while for the other they preferred <v>). To that end, each participant was awarded a consistency score (higher scores indicated more consist spelling patterns) and the entire sample was then divided in two groups through a median split. Surprisingly, the orthographic skeletons effect (i.e., the difference in reading times between trained words with unique and those with unpreferred spellings) was larger in the group of participants who tended to vary in their spelling preferences. This suggests that participants who were less consistent in their preferences for words with two possible spellings are those who were more likely to create orthographic skeletons. One possible explanation of this finding, in line with the Lexical Quality Hypothesis (Perfetti & Hart, 2001), would be that participants who tend to vary in their spelling preferences think more about possible spellings of the novel words they are presented with and how good of a fit a particular grapheme is in a particular word. In other words, these participants generate more solid orthographic skeletons which would lead to larger orthographic skeleton effects as well as better recall of the novel words. By contrast, participants who tend to hold on to one particular letter across all words, might do so as it is less demanding. They may therefore be less prone to think about and imagine the possible spellings of these novel words, and thus generate vague orthographic representations. Consequently, they would not show strong orthographic skeleton effects and would not have good visual cues when recalling the words. However, the observed individual differences in generating orthographic skeletons related to phonological short-term memory (PSTM) skills could also be interpreted the other way around. That is, participants with better PSTM skills are able to create better visual memory traces in the form of preliminary orthographic representations in their episodic memory. This would lead to both better word reading as well as larger orthographic skeleton effects. This interpretation is supported by EEG data showing individual differences in word learning related to reading comprehension skills (Perfetti et al., 2005). In a novel word learning study, Perfetti and colleagues (2005) showed that participants with better reading comprehension skills generated better memory traces for newly acquired words as indicated by their stronger episodic memory effects at around 400-600ms after written word onset. Importantly, they were also better at learning novel words as compared to less skilled comprehenders. These findings thus show that both reading skills and memory capacity are important for novel word acquisition. Future research is however needed in order to explore the origin of the relationship between the two, and specifically, test whether one is a necessary precursor of the other.

Finally, all previous studies had been done with L1 speakers, who were tested in their first and most dominant language. Acquiring oral language skills in one's native language almost always precedes the acquisition of reading and writing. By contrast, in second language(s) (L2), which are acquired later in life, and most often in a classroom context, the two skills are usually developed in parallel. It could therefore be the case that orthographic effects in spoken word learning may actually be stronger in an L2, since oral and written skills attain a more equal status in languages acquired in a classroom setting. This prediction is supported by findings reporting positive effects of orthographic exposure on novel form word learning in an L2 (Bürki et al., 2019). When trained on novel word forms presented either in their auditory form only or along with their spellings, French L1 learners of English showed better learning outcomes (i.e., fewer errors and faster naming times) for words learned in both auditory and orthographic forms. Therefore, orthography as a mnemonic tool could explicitly be used in second language instruction. L2-learners could be prompted to create orthographic representations every time they learn a novel spoken word, despite the possibility of creating an *incorrect* representations, since this would give them a more solid memory cue of a newly acquired word.

Orthographic effects in spoken language processing

Apart from better understanding the link between oral vocabulary knowledge and word reading, the work conducted in the course of this thesis also aimed to expand on the previously reported orthographic effects in spoken word processing. To test whether orthographic knowledge affects spoken language processing, and specifically, speech perception, a vast amount of previous research has made use of sound-to-spelling (in)consistencies and has tested differences in recognition of words with one, as compared to those with multiple possible spellings. Faster recognition times for spoken words with unique spellings obtained in various different languages (French, English, Portuguese and Spanish) as well as populations differing in their reading skills (children and adults), were taken as evidence that, once acquired, orthography indeed changes the perception of spoken words (Chéreau et al., 2007; Pattamadilok et al., 2009; Ventura et al., 2004; Ziegler & Ferrand, 1998).

The three experiments presented here, along with the work by Wegener and colleagues (Wegener et al., 2018, 2020), add to this line of research by revealing another way in which orthography can influence speech perception. By teaching participants spoken words with one or two possible spellings, we show that orthographic knowledge can lead to the creation of orthographic representations of novel words never seen in writing. Importantly, data from Experiment 3 suggest that orthographic effects may sometimes be driven by participants' strategies rather than being completely automatic in nature. This contrasts with the previous work on speech perception. Even though they still debate on whether orthographic effects are due to the automatic co-activation of orthographic representation during speech perception (Chéreau et al., 2007), or whether existing phonological representations get modified during reading acquisition (Muneaux & Ziegler, 2004; Perre et al., 2009), previous studies showing orthographic effects in speech perception agree that the observed effects are not under participants' control. Replicating the pattern of results from Experiment 1 only in the group of active learners from Experiment 3, that is, learners who were aware of the learning process, suggests that the orthographic effects seen in spoken word learning could actually be strategic. This findings implies that the bi-directional links between phonological and orthographic representations may under some circumstances be sensitive to top-down influences.

Future directions

The three experiments presented here were carefully planned and designed to test three very specific research questions. Consequently, we tried to control for all confounding variables that could potentially mask the effects we were interested in. That being said, it is important to outline the limitations of the present work with the aim to motivate future research and hence broaden our knowledge on the link between orthography and auditory word learning.

Firstly, all our conclusions are based on the assumption that the reported effects of aural training on the subsequent word reading emerge because participants generated preliminary orthographic representations during the learning phase. This thus implies that the newly acquired orthographic representations must have been kept in lexical memory, along with the already existing representations of familiar words. Nevertheless, another equally possible explanation of the presented results is in line with the aforementioned accounts postulating that the effects of aural training on word reading are taking place *at the moment* of the first visual encounter. According to this alternative explanation, orthographic representations are not generated during the learning process and then retained in the mental lexicon, but are rather generated at the first visual encounter with novel words' spellings. Specifically, it could be the case that when first confronted with the written form of an aurally familiar word (e.g., <vadi>), a phonological representation of this word is automatically generated. This phonological representation would match the one generated during aural training, and at the same time activate its preferred orthographic form. The latter, being in line with the actual spelling, would facilitate word recognition. By contrast, the newly created orthographic representation might not overlap with the spelling leading to its slower recognition. Having said that, it is important to note that the predictions derived from this alternative hypothesis do not differ from those implied by the orthographic skeleton hypothesis. This makes it difficult, if not impossible, to adjudicate between the two given the data we have obtained so far. In order to show which of the two mechanisms (if not both at the same time), is more likely to be taking place during the first visual encounter with aurally familiar words, future research should turn to more precise and online measure of language processing such as EEG or eye-tracking. In addition, formulating computational models which are able to simulate the observed data (see below) would help better describe the underlying process driving the observed orthographic effects.

Another important consideration to take into account is the level of reading skills needed to generate orthographic skeletons, and in particular, to generate them for all aurally acquired words regardless of the number of possible spellings. The data we present here may be limited to skilled readers. Due to their vast experience with written language (in the form of reading and writing), skilled readers may be impacted by orthography more than children. Consequently, they might be more prone to generating orthographic skeletons for all novel words, even for those whose spellings were uncertain. By contrast, early and developing readers lack the reading expertise and experience with orthotactic probabilities present in skilled readers. As a result, they may generate orthographic skeletons only when a single highly probable spelling is available (i.e., only for words with unique spellings). Future research could set out to test children at different stages of reading acquisition in order to provide more insight into the developmental trajectory of the mechanism responsible for generating orthographic skeletons. This way we would be able to determine the extent to which the orthographic skeleton hypothesis can be generalized to different populations (i.e., early readers).

In addition, the findings we observed may also be specific to readers who successfully acquired sound-to-spellings correspondences (i.e., normally-developing readers). Since generating orthographic representations requires good phonological and orthographic processing skills, readers with dyslexia, who have difficulties with both, may actually not be able to generate orthographic skeletons as a result of aural training. Indeed, previous research has shown that both phonological and orthographic representations tend to be deficient in this population (Cao, Bitan, Chou, Burman, & Booth, 2006; Elbro & Jensen, 2005; Hasko, Bruder, Bartling, & Schulte-Körne, 2012). Moreover, children with dyslexia usually need more time to create associations between novel phonological representations and novel semantic concepts (Elbro & Jensen, 2005). Generating orthographic skeletons may thus require more extensive training with spoken words in readers with reading difficulties such as dyslexia. However, since this group particularly struggles with phonological decoding, which consists of mapping orthographic input into already existing phonological representations, extensive aural training with novel phonological word forms may be beneficial. Generating any kind of prior orthographic representations would help decoding the words when they are first encountered in print. Therefore, adapting the tasks used in the present thesis so they could be used to test readers with dyslexia could result in the creation of new reading interventions.

Finally, although we tried to describe the mechanism(s) by which orthographic skeletons are generated during spoken word learning as thoroughly as possible, experimental work alone is limited when it comes to building complete theoretical accounts. To better understand the exact processes underlying the creation of orthographic representations during spoken word learning, a formal computational model able to account for the observed data, as well as simulate future outcomes, should be formulated (see Rooij, 2022, for the importance of models in understanding the phenomena studied in experimental psychology). Such a model would have to be complete (describe and define all mechanisms involved in the process of generating orthographic skeletons) and sufficient (be able to account for all the observed data taking into account the properties of each writing system). As a result, the model would be able to generate predictions that would then be tested against participants' behavior, this way advancing the theoretical framework. A first step in this direction could be to try to incorporate the present data into already existing model structures (e.g., the TRACE model; McClelland & Elman, 1986) by tweaking some parameters specific to the studied phenomenon. This would give an overview of the processes which are underspecified and hence need to be redefined.

Conclusion

The present thesis investigated the role of orthography in spoken word learning. To that end, we tested whether the previously observed positive link between oral vocabulary knowledge and word reading could be mediated by a mechanism responsible for converting newly acquired phonological representations into orthographic ones even before the first visual encounter with the words' actual spellings. Across three experiments we show that skilled readers do generate preliminary orthographic representations (so-called orthographic skeletons) for newly acquired spoken words. Importantly, our work provides some of the first evidence that orthographic skeletons are generated even when there is uncertainty regarding the correct spelling, as is the case when multiple spelling options are possible. Moreover, we show that orthographic skeletons are generated in languages with fairly consistent writing systems (e.g., Spanish), but also those with more complex sound-to-spelling correspondences, in which the overall probability of generating an incorrect representation is higher (e.g., French). Finally, our data seem to suggest that generating orthographic skeletons may not be an automatic process inherent to the cognitive system, but could actually be driven by participants' personal strategies. Overall, the present thesis adds to the previous literature by describing new ways in which orthographic knowledge affects spoken language processing.

List of Figures

List of Tables

A Appendices

A.1 Fillers from the Spanish novel word spelling task

Table A.1: Spanish Fillers

Consistent	Inonsistent	
/nufa/	/besu/	/kika/
/lifa/	/biɲo/	/kodu/
/lusi/	/baru/	/kado/
/tado/	/beo/	/kebo/
/fasa/	/bugo/	/kesi/
/dofa/	/beli/	/kibe/
/nadu/	/boi/	/kigo/
/mita/	/bafa/	/keʎo/
/nafo/	/ʎado/	/χibu/
/meli/	/ʎebi/	/χeko/
/teda/	/ʎoto/	/χeɲa/
/tefi/	/ʎubo/	/χeʎa/
/mafe/	/ʎomo/	/χifa/
/lono/	/ʎaku/	/χega/
/puda/	/ʎepa/	/χigo/
/pefo/	/ʎuχa/	/χeru/

A.2 Spellings from the Spanish novel word spelling task

In the pre-test spelling task 5.01% of all inconsistent items (words with two possible spellings) were misspelled due to either leaving out or misspelling the first and hence, target phoneme. In cases where participants left out or misspelled a target sound (e.g., wrote 'chuñe' pronounced as /tʃuɲe/, instead of 'yuñe' or 'lluñe' for the item /ʎuɲe/) their preferred spellings were inferred from the correct spellings of both the filler words and other target items starting with the same syllable. Specifically, for each misspelled item we looked at how all the other words starting with the same syllable were spelled and the syllable used most of the times was chosen as the preferred one.

In both Tables A.8 and A.9 numbers in red indicate novel words whose spellings had to be inferred from the other novel word starting with the same syllable. Numbers between brackets represent the preferred spellings after correcting for misspellings.

Table A.2: Preferred Spanish Spellings from the Set A

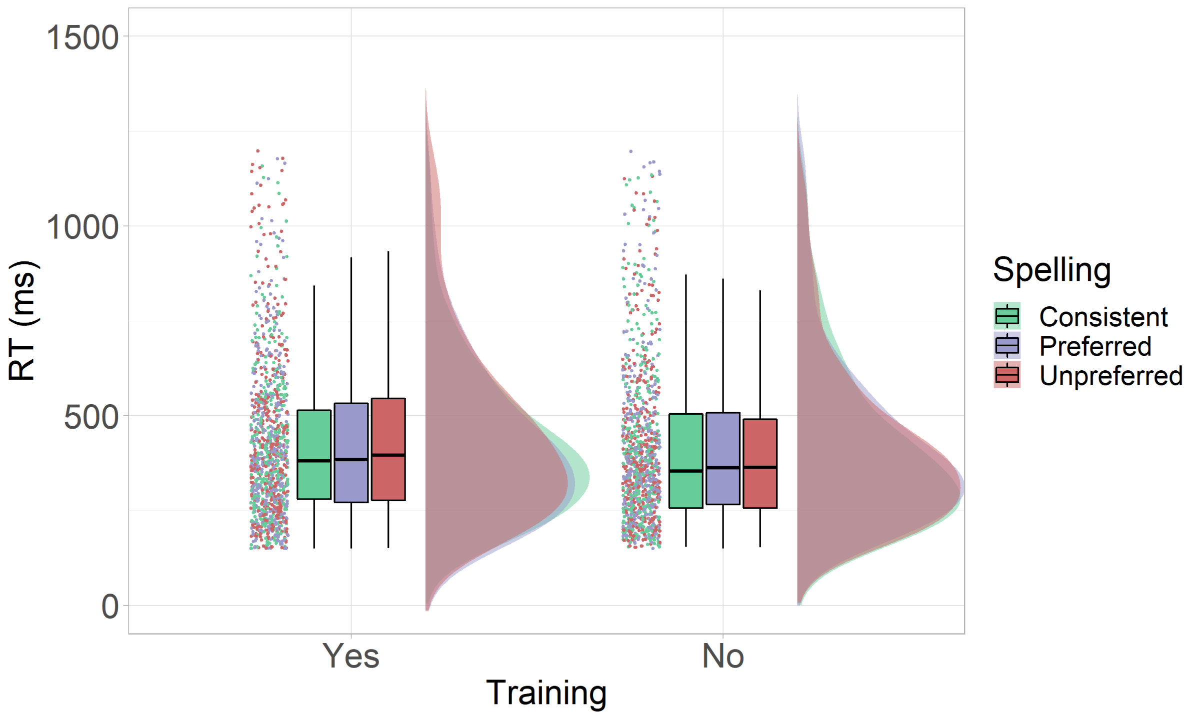
Set A			
	Phonological Form	**Possible Spellings**	**Preferred Spelling Per Participant**
Inconsistent Preferred	/ʎedu/	yedu	36 (39)
		lledu	9 (9)
	/ʎuɲe/	yuñe	22 (41)
		lluñe	5 (7)
	/χepo/	gepo	30
		jepo	18
	/χifo/	gifo	19
		jifo	29
	/bamu/	bamu	20 (20)
		vamu	27 (28)
	/bupe/	bupe	34 (48)
		vupe	0 (0)
	/kime/	kime	14
		quime	34
	/ketu/	ketu	14
		quetu	34
Inconsistent Unpreferred	/ʎefo/	yefo	31 (34)
		llefo	14 (14)
	/ʎupo/	yupo	43
		llupo	5
	/χede/	gede	27
		jede	21
	/χitu/	gitu	21
		jitu	27
	/badi/	badi	40 (41)
		vadi	7 (7)
	/bumi/	bumi	44 (47)
		vumi	1 (1)
	/kifo/	kifo	10 (11)
		quifo	35 (37)
	/keli/	keli	17
		queli	31

Table A.3: Preferred Spanish Spellings from the Set B

	Set B		
	Phonological Form	**Possible Spellings**	**Preferred Spelling Per Participant**
Inconsistent Preferred	/ʎepo/	yepo	36 (38)
		llepo	10 (10)
	/ʎule/	yule	39 (40)
		llule	8 (8)
	/χeni/	geni	34
		jeni	14
	/χipe/	gipe	23 (23)
		jipe	24 (25)
	/bafu/	bafu	40 (46)
		vafu	0 (2)
	/buɲe/	buñe	36 (48)
		vuñe	0 (0)
	/kipe/	kipe	10
		quipe	38
	/kefi/	kefi	14
		quefi	34
Inconsistent Unpreferred	/ʎeli/	yeli	39 (40)
		lleli	8 (8)
	/ʎufi/	yufi	42 (44)
		llufi	4 (4)
	/χetu/	getu	26
		jetu	22
	/χidu/	gidu	26
		jidu	22
	/bani/	bani	35
		vani	13
	/buti/	buti	45 (47)
		vuti	1 (1)
	/kiɲo/	kiño	12
		quiño	36
	/kedi/	kedi	11
		quedi	37

A.3 Distribution of Reaction Times From the Experiment 1

Figure A3: Distribution of Reaction Times From the Experiment 1



Raincloud plots showing the distributions of RTs for both trained and untrained consistent, preferred and unpreferred spellings.

A.4 Fillers from the French novel word spelling task

Table A.4: French Fillers

Consistent	Inonsistent	
/bedavə/	/ʒinymə/	/ʒibybə/
/dabinə/	/ʒetivə/	/ʒebabə/
/tubanə/	/sebanə/	/sevivə/
/mavymə/	/sibyvə/	/sidabə/
/nydivə/	/fatavə/	/fanymə/
/vemagə/	/femidə/	/fetogə/
/pivavə/	/kimybə/	/kilydə/
/lydanə/	/ketagə/	/kemumə/
/bamivə/	/ʒimanə/	/ʒipagə/
/dedabə/	/ʒebanə/	/ʒelyvə/
/talynə/	/simamə/	/simymə/
/medivə/	/selidə/	/sedivə/
/nedumə/	/fabigə/	/fadumə/
/vamyvə/	/femydə/	/fetabə/
/pitogə/	/kinynə/	/kilavə/
/labydə/	/kedibə/	/kepymə/

A.5 French Pseudowords from the Pseudoword Repetition Task

Table A.5: French-like pseudiwords from the pseudoword repetition task

Practice trials	/kɛd/	/ʃis/						
	/lɛp/	/fyb/	/mav/					
2 items sequences	/tɔʃ/	/nyp/						
	/fik/	/ʁœd/						
3 items sequences	/lyz/	/zɑ̃ʃ/	/nɔl/					
	/dyl/	/nɔk/	/zɛt/					
4 items sequences	/suv/	/zɔd/	/ʃym/	/nɛl/				
	/nys/	/tup/	/mab/	/fɛʃ/				
5 items sequences	/sod/	/vɛk/	/lɔ̃f/	/ʒab/	/mip/			
	/pɛm/	/zat/	/nuʁ/	/vɔʃ/	/duk/			
6 items sequences	/dɑ̃ʒ/	/gas/	/miv/	/bɛf/	/nyt/	/ʒyk/		
	/mɔs/	/ʃœʁ/	/ʒov/	/kɛz/	/gun/	/lib/		
7 items sequences	/ʁɛp/	/ʒɔʃ/	/lyv/	/zem/	/foz/	/nœʁ/	/ʒib/	
	/vad/	/pyn/	/loz/	/niʒ/	/bɛʁ/	/kɑ̃ʃ/	/fœt/	
8 items sequences	/kof/	/ʒam/	/tɛn/	/zyʁ/	/puv/	/fœʁ/	/ʁɑ̃b/	/diʃ/
	/daɲ/	/fɛk/	/pɔs/	/nuv/	/zit/	/ʃyd/	/bɔ̃ʒ/	/ʁœl/

A.6 Spellings from the French novel word spelling task

Table A.6: Preferred French Spellings from the Set A

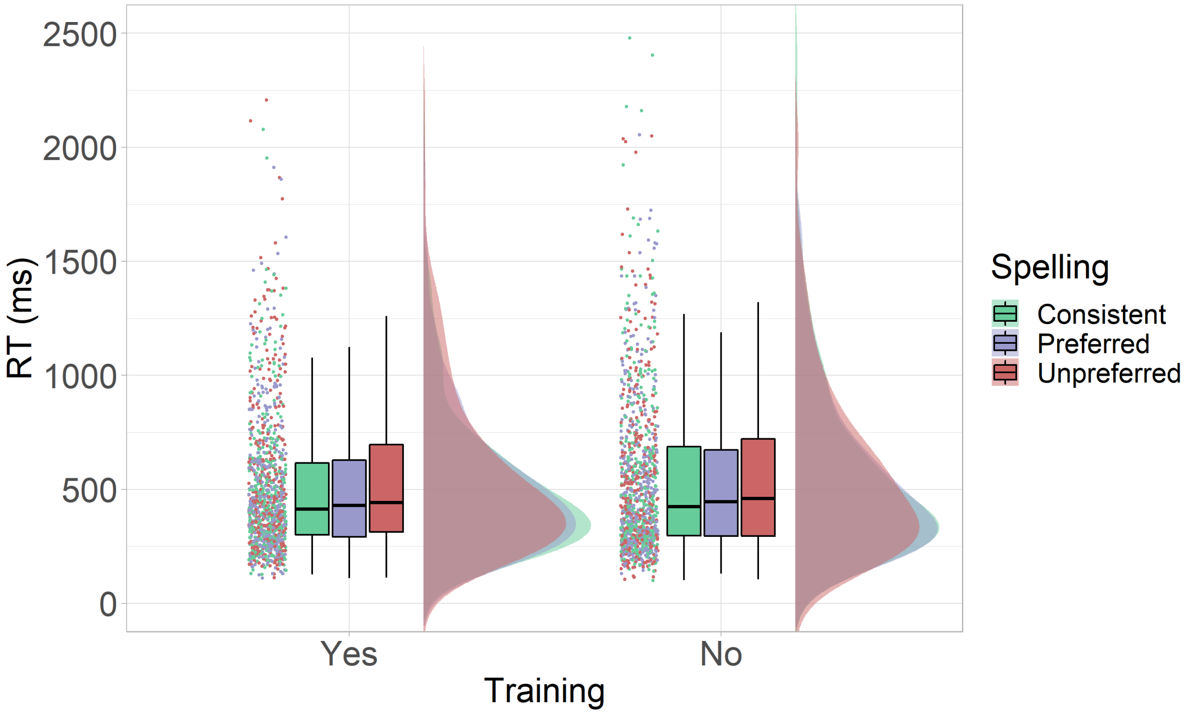
Set A			
	Phonological Form	**Possible Spellings**	**Preferred Spelling Per Participant**
Inconsistent Preferred	/ʒinavə/	ginav	38
		jinav	8
	/ʒebinə/	gébine	43
		jébine	3
	/sedunə/	sédoune	13
		cédoune	33
	/simybə/	simub	26
		cimub	20
	/fapyvə/	fapuve	41
		phapuve	5
	/fedinə/	fédine	41
		phédine	5
	/kityvə/	quituve	23
		kituve	23
	/kemagə/	quémague	22
		kémague	24
Inconsistent Unpreferred	/ʒitymə/	gitume	35
		jitume	11
	/ʒevabə/	gévab	39
		jévab	7
	/semivə/	sémive	21
		cémive	25
	/sibavə/	sibave	24
		cibave	24
	/fanynə/	fanune	41
		phanune	5
	/fenɔgə/	fénogue	35
		phénogue	11
	/kidunə/	quidoune	25
		kidoune	15
	/kepydə/	quépude	25
		képude	15

Table A.7: Preferred French Spellings from the Set B

	Set B		
	Phonological Form	**Possible Spellings**	**Preferred Spelling Per Participant**
Inconsistent Preferred	/ʒedavə/	gédave	38
		jédave	8
	/ʒimunə/	gimoune	31
		jimoune	15
	/sepidə/	sépide	19
		cépide	27
	/sitavə/	sitave	17
		citave	29
	/fabɔgə/	fabogue	41
		phabogue	5
	/fenybə/	fénube	37
		phénube	9
	/kenivə/	quénive	22
		kénive	24
	/kipynə/	quipune	25
		kipune	15
Inconsistent Unpreferred	/ʒenyvə/	génuve	40
		jénuve	6
	/ʒitɔgə/	gitogue	34
		jitogue	12
	/sebavə/	sébave	20
		cébave	26
	/sidynə/	sidune	27
		cidune	19
	/fapunə/	fapoune	40
		phapoune	6
	/febadeə/	fébade	39
		phébade	7
	/kepanə/	quépane	19
		képane	27
	/kimavə/	quimave	28
		kimave	18

A.7 Distribution of Reaction Times From the Experiment 2

Figure A7: Distribution of Reaction Times From the Experiment 2



Raincloud plots showing the distributions of RTs for both trained and untrained consistent, preferred and unpreferred spellings.

A.8 Spellings from the Spanish novel word spelling task: Experiment 3

Table A.8: Preferred Spanish Spellings from the Set A

	Set A		
	Phonological Form	**Possible Spellings**	**Preferred Spelling Per Participant**
Inconsistent Preferred	/ʎedu/	yedu	38
		lledu	16
	/ʎuɲe/	yuñe	49
		lluñe	7
	/χepo/	gepo	30
		jepo	32
	/χifo/	gifo	17
		jifo	43
	/bamu/	bamu	25
		vamu	38
	/bupe/	bupe	62
		vupe	2
	/kime/	kime	20
		quime	43
	/ketu/	ketu	20
		quetu	41
Inconsistent Unpreferred	/ʎefo/	yefo	32
		llefo	20
	/ʎupo/	yupo	48
		llupo	5
	/χede/	gede	30
		jede	33
	/χitu/	gitu	25
		jitu	38
	/badi/	badi	48
		vadi	16
	/bumi/	bumi	61
		vumi	2
	/kifo/	kifo	5
		quifo	37
	/keli/	keli	27
		queli	34

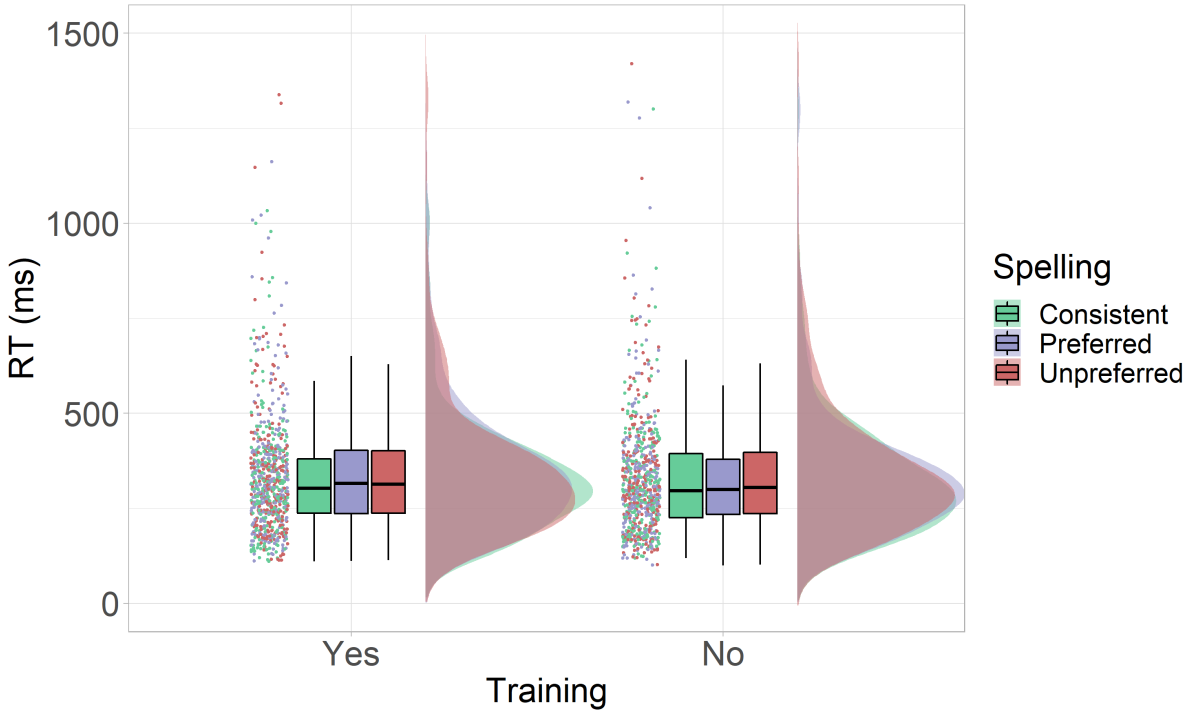
Table A.9: Preferred Spanish Spellings from the Set B

	Set B		
	Phonological Form	**Possible Spellings**	**Preferred Spelling Per Participant**
Inconsistent Preferred	/ʎepo/	yepo	42
		llepo	10
	/ʎule/	yule	50
		llule	6
	/χeni/	geni	47
		jeni	16
	/χipe/	gipe	27
		jipe	35
	/bafu/	bafu	58
		vafu	6
	/buɲe/	buñe	63
		vuñe	0
	/kipe/	kipe	22
		quipe	42
	/kefi/	kefi	26
		quefi	37
Inconsistent Unpreferred	/ʎeli/	yeli	40
		lleli	11
	/ʎufi/	yufi	48
		llufi	7
	/χetu/	getu	24
		jetu	40
	/χidu/	gidu	29
		jidu	34
	/bani/	bani	37
		vani	27
	/buti/	buti	64
		vuti	0
	/kiɲo/	kiño	17
		quiño	47
	/kedi/	kedi	18
		quedi	46

A.9 Distribution of Reaction Times From the Experiment 3

Active Learners

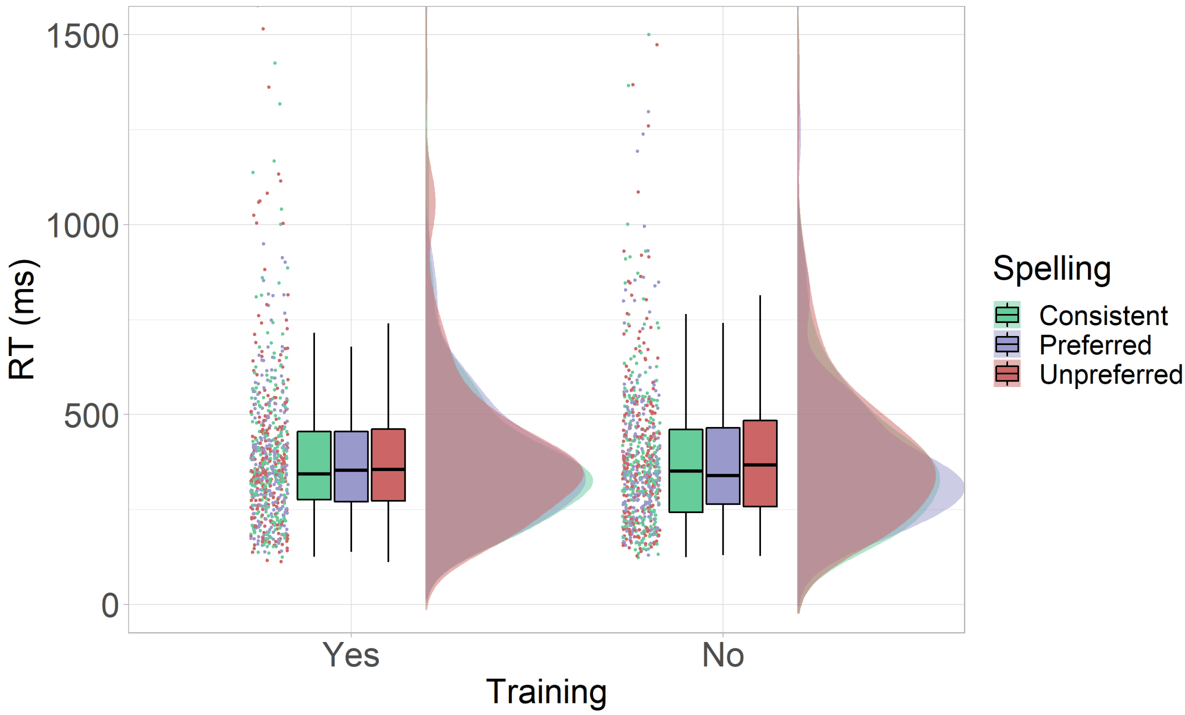
Figure A9.1: Distribution of Reaction Times From the Experiment 3



Raincloud plots showing the distributions of RTs for both trained and untrained consistent, preferred and unpreferred spellings.

Passive Learners

Figure A9.2: Distribution of Reaction Times From the Experiment 3



Raincloud plots showing the distributions of RTs for both trained and untrained consistent, preferred and unpreferred spellings.

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.

Álvarez-Cañizo, M., Suárez-Coalla, P., & Cuetos, F. (2019). Orthographic learning in spanish children: Influence of previous semantic and phonological knowledge. *Journal of Research in Reading*, *42*(1), 137–149.

Anglin, J. M., Miller, G. A., & Wakefield, P. C. (1993). Vocabulary development: A morphological analysis. *Monographs of the Society for Research in Child Development*, i–186.

Baayen, R. (2008). Analyzing linguistic data: A practical introduction to statistics using r. Cambridge university press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412.

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research*, *3*(2), 12–28.

Baddeley, A., Papagno, C., & Vallar, G. (1988). When long-term learning depends on short-term storage. *Journal of Memory and Language*, *27*(5), 586–595.

Baron, L. S., Hogan, T. P., Alt, M., Gray, S., Cabbage, K. L., Green, S., & Cowan, N. (2018). Children with dyslexia benefit from orthographic facilitation during spoken word learning. *Journal of Speech, Language, and Hearing Research*, *61*(8), 2002–2014.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *arXiv Preprint arXiv:1506.04967*.

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

Batterink, L., & Neville, H. (2011). Implicit and explicit mechanisms of word learning in a narrative context: An event-related potential study. *Journal of Cognitive Neuroscience*, *23*(11), 3181–3196.

Berent, I., & Perfetti, C. A. (1995). A rose is a REEZ: The two-cycles model of phonology assembly in reading english. *Psychological Review*, *102*(1), 146.

Beyersmann, E., Wegener, S., Nation, K., Prokupzcuk, A., Wang, H.-C., & Castles, A. (2021). Learning morphologically complex spoken words: Orthographic expectations of embedded stems are formed prior to print exposure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *47*(1), 87.

Bloomfield, L. (1933). *Language*. New York: Holt, Rinehart, & Winston.

Box, G. E., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, *26*(2), 211–243.

Brysbaert, M. (2022). Word recognition II: Phonological coding in reading. *The Science of Reading: A Handbook*, 79–101.

Brysbaert, M., Stevens, M., Mandera, P., & Keuleers, E. (2016). How many words do we know? Practical estimates of vocabulary size dependent on word definition, the degree of language input and the participant's age. *Frontiers in Psychology*, *7*, 1116.

Brysbaert, M., Van Dyck, G., & Van de Poel, M. (1999). Visual word recognition in bilinguals: Evidence from masked phonological priming. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(1), 137.

Bürki, A., Welby, P., Clément, M., & Spinelli, E. (2019). Orthography and second language word learning: Moving beyond "friend or foe?" *The Journal of the Acoustical Society of America*, *145*(4), EL265–EL271.

Burnham, D. (2003). Language specific speech perception and the onset of reading. *Reading and Writing*, *16*(6), 573–609.

Cao, F., Bitan, T., Chou, T.-L., Burman, D. D., & Booth, J. R. (2006). Deficient orthographic and phonological representations in children with dyslexia revealed by brain activation patterns. *Journal of Child Psychology and Psychiatry*, *47*(10), 1041–1050.

Carrillo, M. S., & Alegría, J. (2014). The development of children's sensitivity to bigram frequencies when spelling in spanish, a transparent writing system. *Reading and Writing*, *27*(3), 571–590.

Castles, A., & Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read? *Cognition*, *91*(1), 77–111.

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

Castles, A., & Nation, K. (2008). Learning to be a good orthographic reader. *Journal of Research in Reading*, *31*(1), 1–7.

Chéreau, C., Gaskell, M. G., & Dumay, N. (2007). Reading spoken words: Orthographic effects in auditory priming. *Cognition*, *102*(3), 341–360.

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.

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.

Daneman, M., & Stainton, M. (1991). Phonological recoding in silent reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(4), 618.

Davis, C. J., & Perea, M. (2005). BuscaPalabras: A program for deriving orthographic and phonological neighborhood statistics and other psycholinguistic indices in spanish. *Behavior Research Methods*, *37*(4), 665–671.

De Bruin, A., Carreiras, M., & Duñabeitia, J. A. (2017). The BEST dataset of language proficiency. *Frontiers in Psychology*, *8*, 522.

De Saussure, F. (1989). *Cours de linguistique générale* (Vol. 1). Otto Harrassowitz Verlag.

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.

Dickens, C. (1949). *A tale of two cities [1859]*. Gawthorn.

Dijkstra, T., Roelofs, A., & Fieuws, S. (1995). Orthographic effects on phoneme monitoring. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *49*(2), 264.

Donnenwerth-Nolan, S., Tanenhaus, M. K., & Seidenberg, M. S. (1981). Multiple code activation in word recognition: Evidence from rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory*, *7*(3), 170.

Downey, S. S., Sun, G., & Norquest, P. (2017). alineR: an R Package for Optimizing Feature-Weighted Alignments and Linguistic Distances. *The R Journal*, *9*(1), 138–152. http://doi.org/10.32614/RJ-2017-005

Drewnowski, A., & Healy, A. F. (1982). Phonetic factors in letter detection: A reevaluation. *Memory & Cognition*, *10*(2), 145–154.

Duyck, W., Diependaele, K., Drieghe, D., & Brysbaert, M. (2004). The size of the cross-lingua masked phonological priming effect does not depend on second language proficiency. *Experimental Psychology*, *51*(2), 116.

Ehri, L. C. (1978). Beginning reading from a psycholinguistic perspective: Amalgamation of word identities. *The Development of the Reading Process*, 1–33.

Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading*, *9*(2), 167–188.

Ehri, L. C. (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Scientific Studies of Reading*, *18*(1), 5–21.

Ehri, L. C., & Wilce, L. S. (1979). The mnemonic value of orthography among beginning readers. *Journal of Educational Psychology*, *71*(1), 26.

Elbro, C., & Jensen, M. N. (2005). Quality of phonological representations, verbal learning, and phoneme awareness in dyslexic and normal readers. *Scandinavian Journal of Psychology*, *46*(4), 375–384.

Elbro, C., & Pallesen, B. R. (2002). The quality of phonological representations and phonological awareness: A causal link. *Precursors of Functional Literacy*, *11*, 17–32.

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(4), 680.

Friesen, D. C., & Jared, D. (2012). Cross-language phonological activation of meaning: Evidence from category verification. *Bilingualism: Language and Cognition*, *15*(1), 145–156.

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics*, *27*(4), 513–543.

Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language*, *28*(2), 200–213.

Gerbier, E., & Toppino, T. C. (2015). The effect of distributed practice: Neuroscience, cognition, and education. *Trends in Neuroscience and Education*, *4*(3), 49–59.

Goff, D. A., Pratt, C., & Ong, B. (2005). The relations between children's reading comprehension, working memory, language skills and components of reading decoding in a normal sample. *Reading and Writing*, *18*(7), 583–616.

Hasko, S., Bruder, J., Bartling, J., & Schulte-Körne, G. (2012). N300 indexes deficient integration of orthographic and phonological representations in children with dyslexia. *Neuropsychologia*, *50*(5), 640–654.

Hoenig, J. M., & Heisey, D. M. (2001). The abuse of power: The pervasive fallacy of power calculations for data analysis. *The American Statistician*, *55*(1), 19–24.

Hogaboam, T. W., & Perfetti, C. A. (1978). Reading skill and the role of verbal experience in decoding. *Journal of Educational Psychology*, *70*(5), 717.

Hoonhorst, I., Colin, C., Markessis, E., Radeau, M., Deltenre, P., & Serniclaes, W. (2009). French native speakers in the making: From language-general to language-specific voicing boundaries. *Journal of Experimental Child Psychology*, *104*(4), 353–366.

Hoonhorst, I., Medina, V., Colin, C., Markessis, E., Radeau, M., Deltenre, P., & Serniclaes, W. (2011). Categorical perception of voicing, colors and facial expressions: A developmental study. *Speech Communication*, *53*(3), 417–430.

Horst, J. S., & Hout, M. C. (2016). The novel object and unusual name (NOUN) database: A collection of novel images for use in experimental research. *Behavior Research Methods*, *48*(4), 1393–1409.

Huey, E. B. (1908). *The psychology and pedagogy of reading.* Macmillan.

Hugo, V. (1862). *Les misérables: L'idylle rue plumet et l'épopée rue saint-denis.* C. Lassalle.

Hulme, C., & Snowling, M. J. (2013). Learning to read: What we know and what we need to understand better. *Child Development Perspectives*, *7*(1), 1–5.

Izura, C., Cuetos, F., & Brysbaert, M. (2014). Lextale-esp: A test to rapidly and efficiently assess the spanish vocabulary size. *Psicológica*, *35*(1), 49–66.

Jevtović, M., Stoehr, A., Klimovich-Gray, A., Antzaka, A., & Martin, C. D. (2022). One-to-one or one too many? Linking sound-to-letter mappings to speech sound perception and production in early readers. *Journal of Speech, Language, and Hearing Research*, 1–13.

Johnston, M., Castles, A., et al. (2003). Dissociating automatic orthographic and phonological codes in lexical access and lexical acquisition. *Masked Priming: The State of the Art*, 193–222.

Johnston, Michael, McKague, M., & Pratt, C. (2004). Evidence for an automatic orthographic code in the processing of visually novel word forms. *Language and Cognitive Processes*, *19*(2), 273–317.

Kolinsky, R., Navas, A. L., Paula, F. V. de, Brito, N. R. de, Medeiros Botecchia, L. de, Bouton, S., & Serniclaes, W. (2021). The impact of alphabetic literacy on the perception of speech sounds. *Cognition*, *213*, 104687.

Kondrak, G. (1999). Alignment of phonetic sequences. *Department of Computer Science, University of Toronto, Tech. Rep. CSRG-402*.

Konopak, B., Sheard, C., Longman, D., Lyman, B., Slaton, E., Atkinson, R., & Thames, D. (1987). Incidental versus intentional word learning from context. *Reading Psychology: An International Quarterly*, *8*(1), 7–21.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*, 1–26.

Lange, K., Kühn, S., & Filevich, E. (2015). " just another tool for online studies”(JATOS): An easy solution for setup and management of web servers supporting online studies. *PloS One*, *10*(6), e0130834.

Liuzza, R. M. (1996). Orthography and historical linguistics. *Journal of English Linguistics*, *24*(1), 25–44.

Lukatela, G., & Turvey, M. (2000). Do spelling variations affect associative and phonological priming by pseudohomophones? *Perception & Psychophysics*, *62*(1), 196–217.

Lundberg, I., Frost, J., & Peterson, O. (1991). Phonemic awareness can be developed without reading instruction. *Phonological Processes in Literacy: A Tribute to Isabelle Y. Liberman*, 47–53.

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.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing type i error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*(1), 1–86.

McKague, M., Davis, C., Pratt, C., & Johnston, M. B. (2008). The role of feedback from phonology to orthography in orthographic learning: An extension of item-based accounts. *Journal of Research in Reading*, *31*(1), 55–76.

McKague, M., Pratt, C., & Johnston, M. B. (2001). The effect of oral vocabulary on reading visually novel words: A comparison of the dual-route-cascaded and triangle frameworks. *Cognition*, *80*(3), 231–262.

McMurray, B., Kapnoula, E. C., & Gaskell, M. G. (2016). Learning and integration of new word-forms: Consolidation, pruning, and the emergence of automaticity. In *Speech perception and spoken word recognition* (pp. 126–152). Psychology Press.

Miles, K. P., Ehri, L. C., & Lauterbach, M. D. (2016). Mnemonic value of orthography for vocabulary learning in monolinguals and language minority english-speaking college students. *Journal of College Reading and Learning*, *46*(2), 99–112.

Miller, K. M., & Swick, D. (2003). Orthography influences the perception of speech in alexic patients. *Journal of Cognitive Neuroscience*, *15*(7), 981–990.

Morais, J., Cary, L., Alegria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition*, *7*(4), 323–331.

Morais, J., & Kolinsky, R. (2017). The consequences of phonemic awareness. In *Speech and reading* (pp. 317–337). Routledge.

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

Nation, K., & Snowling, M. J. (2004). Beyond phonological skills: Broader language skills contribute to the development of reading. *Journal of Research in Reading*, *27*(4), 342–356.

Nation, P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. *Vocabulary: Description, Acquisition and Pedagogy*, *14*, 6–19.

Nelson, J. R., Balass, M., & Perfetti, C. A. (2005). Differences between written and spoken input in learning new words. *Written Language & Literacy*, *8*(2), 25–44.

Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology*, *98*(3), 554.

Ouellette, G. P., & Beers, A. (2010). A not-so-simple view of reading: How oral vocabulary and visual-word recognition complicate the story. *Reading and Writing*, *23*(2), 189–208.

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.

Perfetti, C. A., Bell, L. C., & Delaney, S. M. (1988). Automatic (prelexical) phonetic activation in silent word reading: Evidence from backward masking. *Journal of Memory and Language*, *27*(1), 59–70.

Perfetti, C. A., & Hart, L. (2001). The lexical basis of comprehension skill.

Perfetti, C. A., Wlotko, E. W., & Hart, L. A. (2005). Word learning and individual differences in word learning reflected in event-related potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(6), 1281.

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.

Perre, L., & Ziegler, J. C. (2008). On-line activation of orthography in spoken word recognition. *Brain Research*, *1188*, 132–138.

Pexman, P. M., Lupker, S. J., & Reggin, L. D. (2002). Phonological effects in visual word recognition: Investigating the impact of feedback activation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(3), 572.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, *103*, 56–115.

R Core Team. (2021). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Rabe, M. M., Kliegl, R., & Schad, D. J. (2021). *Designr: Balanced factorial designs.* Retrieved from https://CRAN.R-project.org/package=designr

Rack, J., Hulme, C., Snowling, M., & Wightman, J. (1994). The role of phonology in young children learning to read words: The direct-mapping hypothesis. *Journal of Experimental Child Psychology*, *57*(1), 42–71.

Rastle, K. (2019). EPS mid-career prize lecture 2017: Writing systems, reading, and language. *Quarterly Journal of Experimental Psychology*, *72*(4), 677–692.

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.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, *114*(3), 510.

Raven, J. et al. (2003). Raven progressive matrices. In *Handbook of nonverbal assessment* (pp. 223–237). Springer.

Reber, A. S., & Winter, B. (1994). What manner of mind is this? *Behavioral and Brain Sciences*, *17*(3), 418–419.

Ricketts, J., Bishop, D. V., & Nation, K. (2009). Orthographic facilitation in oral vocabulary acquisition. *Quarterly Journal of Experimental Psychology*, *62*(10), 1948–1966.

Ricketts, J., Nation, K., & Bishop, D. V. (2007). Vocabulary is important for some, but not all reading skills. *Scientific Studies of Reading*, *11*(3), 235–257.

Rooij, I. van. (2022). Psychological models and their distractors. *Nature Reviews Psychology*, *1*(3), 127–128.

Rosenthal, J., & Ehri, L. C. (2008). The mnemonic value of orthography for vocabulary learning. *Journal of Educational Psychology*, *100*(1), 175.

Rozin, P., & Gleitman, L. R. (1977). The structure and acquisition of reading II: The reading process and the acquisition of the alphabetic principle. *Toward a Psychology of Reading*, *66*.

Rutkowska, H. (2017). Orthography. *Historical Outlines from Sound to Text*, 200–217.

Saksida, A., Iannuzzi, S., Bogliotti, C., Chaix, Y., Démonet, J.-F., Bricout, L., et al.others. (2016). Phonological skills, visual attention span, and visual stress in developmental dyslexia. *Developmental Psychology*, *52*(10), 1503.

Salins, A., Leigh, G., Cupples, L., & Castles, A. (2021). Orthographic facilitation of oral vocabulary acquisition in children with hearing loss. *Journal of Speech, Language, and Hearing Research*, *64*(8), 3127–3139.

Sapir, E. (1921). Language: An introduction to the study of speech.

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.

Schimke, E. A., Angwin, A. J., Cheng, B. B., & Copland, D. A. (2021). The effect of sleep on novel word learning in healthy adults: A systematic review and meta-analysis. *Psychonomic Bulletin & Review*, 1–28.

Schmalz, X., Marinus, E., Coltheart, M., & Castles, A. (2015). Getting to the bottom of orthographic depth. *Psychonomic Bulletin & Review*, *22*(6), 1614–1629.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*(4), 523.

Seidenberg, M. S., & Tanenhaus, M. K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory*, *5*(6), 546.

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, *55*(2), 151–218.

Share, D. L. (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, *72*(2), 95–129.

Share, D. L. (2004). Orthographic learning at a glance: On the time course and developmental onset of self-teaching. *Journal of Experimental Child Psychology*, *87*(4), 267–298.

Share, D. L. (2008). Orthographic learning, phonological recoding, and self-teaching. In *Advances in child development and behavior* (Vol. 36, pp. 31–82). Elsevier.

Simon, J. R. (1969). Reactions toward the source of stimulation. *Journal of Experimental Psychology*, *81*(1), 174.

Snowling, M. J. (1995). Phonological processing and developmental dyslexia. *Journal of Research in Reading*.

Snowling, M. J., & Hulme, C. (2021). Annual research review: Reading disorders revisited–the critical importance of oral language. *Journal of Child Psychology and Psychiatry*, *62*(5), 635–653.

Sobczak, J. M., & Gaskell, M. G. (2019). Implicit versus explicit mechanisms of vocabulary learning and consolidation. *Journal of Memory and Language*, *106*, 1–17.

Stuart, M., & Coltheart, M. (1988). Does reading develop in a sequence of stages? *Cognition*, *30*(2), 139–181.

Swank, L. K., & Catts, H. W. (1994). Phonological awareness and written word decoding. *Language, Speech, and Hearing Services in Schools*, *25*(1), 9–14.

Taylor, J., Plunkett, K., & Nation, K. (2011). The influence of consistency, frequency, and semantics on learning to read: An artificial orthography paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(1), 60.

Treiman, R., & Cassar, M. (1997). Can children and adults focus on sound as opposed to spelling in a phoneme counting task? *Developmental Psychology*, *33*(5), 771.

Turan, F., & Guel, G. (2008). Early precursor of reading: Acquisition of phonological awareness skills. *Kuram Ve Uygulamada Egitim Bilimleri*, *8*(1), 279.

Tyler, M. D., & Burnham, D. K. (2006). Orthographic influences on phoneme deletion response times. *Quarterly Journal of Experimental Psychology*, *59*(11), 2010–2031.

Uittenhove, K., Jeanneret, S., & Vergauwe, E. (2022). From lab-based to web-based behavioural research: Who you test is more important than how you test. PsyArXiv. http://doi.org/10.31234/osf.io/uy4kb

Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory & Cognition*, *15*(3), 181–198.

Van Orden, G. C., & Goldinger, S. D. (1994). Interdependence of form and function in cognitive systems explains perception of printed words. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(6), 1269.

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.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, *101*(2), 192.

Wegener, S., Beyersmann, E., Wang, H.-C., & Castles, A. (2022). Oral vocabulary knowledge and learning to read new words: A theoretical review. *Australian Journal of Learning Difficulties*, 1–26.

Wegener, S., Wang, H.-C., Lissa, P. de, Robidoux, S., Nation, K., & Castles, A. (2018). Children reading spoken words: Interactions between vocabulary and orthographic expectancy. *Developmental Science*, *21*(3), e12577.

Wegener, S., Wang, H.-C., Nation, K., & Castles, A. (2020). Tracking the evolution of orthographic expectancies over building visual experience. *Journal of Experimental Child Psychology*, *199*, 104912.

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. Retrieved from https://ggplot2.tidyverse.org

Woods, K. J., Siegel, M. H., Traer, J., & McDermott, J. H. (2017). Headphone screening to facilitate web-based auditory experiments. *Attention, Perception, & Psychophysics*, *79*(7), 2064–2072.

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

Ziegler, J. C., Jacobs, A. M., & Stone, G. O. (1996). Statistical analysis of the bidirectional inconsistency of spelling and sound in french. *Behavior Research Methods, Instruments & Computers*, *28*(4), 504–515.

Ziegler, J. C., & Muneaux, M. (2007). Orthographic facilitation and phonological inhibition in spoken word recognition: A developmental study. *Psychonomic Bulletin & Review*, *14*(1), 75–80.

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.

Ziegler, J. C., Stone, G. O., & Jacobs, A. M. (1997). What is the pronunciation for-ough and the spelling for/u/? A database for computing feedforward and feedback consistency in english. *Behavior Research Methods, Instruments, & Computers*, *29*(4), 600–618.