Spoken L2 words activate L1 orthographic information in late L2 learners

Outi Veivo[1], Vincent Porretta[2], Jukka Hyönä[3] & Juhani Järvikivi[4]

[1]Department of French, School of Languages and Translation Studies, University of Turku, Finland
[2]Department of Psychology, University of Windsor, Ontario, Canada
[3]Department of Psychology, University of Turku, Finland
[4]Department of Linguistics, University of Alberta, Edmonton, Canada

Submitted as: *Research article*

Address for correspondence:
Outi Veivo
French Translation Studies
Koskenniemenkatu 4
20014 University of Turku
Finland
Email: outi.veivo@utu.fi
Tel. +358-44-0402248

Abstract

This study investigated the time-course of activation of orthographic information in spoken word recognition with two visual world eye-tracking experiments in a task where L2 spoken word forms had to be matched with their printed referents. Participants (n = 64) were L1 Finnish learners of L2 French ranging from beginners to highly proficient. In Exp. 1, L2 targets (e.g. *<cidre>* /sidʀ/) were presented with either orthographically overlapping onset competitors (e.g. *<cintre>* /sɛ̃tʀ/) or phonologically overlapping onset competitors (*<cycle>* /sikl/). In Exp. 2, L2 targets (e.g., *<paume>* /pom/) were associated with L1 competitors in conditions symmetric to Exp. 1 (*<pauhu>* /pauhu/ vs. *<pommi>* /pom:i/). In the within-language experiment (Exp. 1), the difference in target identification between the experimental conditions was not significant. In the between-language experiment (Exp. 2), orthographic information impacted the mapping more in lower proficiency learners, and this effect was observed 600ms after the target word onset. The influence of proficiency on the matching was non-linear: proficiency impacted the mapping significantly more in the lower half of the proficiency scale in both experiments. These results are discussed in terms of co-activation of orthographic and phonological information in L2 spoken word recognition.

Introduction

There is a growing body of literature showing that just as phonological information is activated during the processing of written language (see e.g. Frost, 1998, for a review), orthographic information is activated during the processing of spoken language (see Frost & Ziegler, 2007, for a review). There are, however, fewer studies on the role of orthography in second language (L2) spoken word processing. In this study, we are interested in how L2 learners with a formal instruction background use orthographic information in spoken word recognition. Late L2 learners differ from native language (L1) speakers because they already use one phonological system that can influence the learning of another system (e.g. Best & Tyler, 2007), and because they are already familiar with the grapheme-phoneme correspondences of their L1 which can have a strong impact on the perception and learning of L2 sounds (Bassetti, 2006; Escudero, Hayes-Harb, & Mitterer, 2008; Escudero & Wanrooij, 2010; Hayes-Harb, Nicol, & Barker, 2010; Showalter & Hayes-Harb, 2015). Also, unlike L1 speakers who learn orthographic word forms only after the phonological forms have been established, literate L2 learners in formal instruction are exposed to written word forms early on in the learning process. The present study investigated how these L2 learners map L2 spoken words onto their written counterparts; specifically, the extent to which this mapping is mediated by orthographic or phonological information, and to which L1 grapheme-phoneme correspondences are activated in this process. For this purpose, we conducted two experiments

where spoken word forms had to be matched with their printed referents while participants' eye movements were monitored. Importantly, we also set out to evaluate the role of L2 proficiency in this matching process.

Even though orthographic information is not necessary in L1 spoken language processing, it is known to be activated even during (non-metaphonological) language processing tasks like lexical decision that do not demand a special focus on the phonotactic or orthotactic structure of the word forms (Grainger & Ferrand, 1996; Grainger, Diependaele, Spinelli, Ferrand, & Farioli, 2003; Salverda & Tanenhaus, 2010; Spinelli, Ferrand, & Farioli, 2003; Ventura, Morais, Pattamadilok, & Kolinsky, 2004; Ziegler & Ferrand, 1998, but see e.g. Reinisch & Mitterer, 2015 for the lack of orthographic effects in the perception of conversational speech). These orthographic effects have been explained by a simultaneous co-activation of phonological and orthographic representations (e.g. Grainger et al., 2003), or by an activation of orthographically restructured phonological representations (Taft, Castles, Davis, Lazendic, & Nguyen-Hoan, 2008) during the processing of spoken words. In the L1, the written forms of words are learned after their spoken forms, but in L2 instructed learning environments the two modalities are learned in parallel. As a result of this co-structuration of orthographic and phonological information (Veivo & Järvikivi, 2013), orthography may have a more important role in the L2 lexicon than in the L1 lexicon. Further, there is evidence that if the orthographic system of the L2 is incongruent – i.e. if the phonemes can be represented by several different graphemes or vice versa – parallel

acquisition of orthography and phonology can be a hindrance to the acquisition of the L2 phonological system (Escudero, Simon & Mulak, 2014).

In L2 spoken word processing, the activation of orthographic information has been studied especially from the point of view of the parallel activation of the L1. For example, Bartolotti, Daniel, and Marian (2013) showed that during spoken word recognition in a newly acquired L2, orthographically similar L1 word forms are activated even if they are pronounced differently from the target words. This result is complementary to studies showing that phonologically similar words of both languages of bilingual or second language speakers compete for recognition in parallel (Blumenfeld & Marian, 2007; Marian & Spivey, 2003a; Marian & Spivey, 2003b; Spivey and Marian, 1999).

The role of orthographic input for the learning of L2 phonology has been widely studied (for reviews, see Bassetti, 2008; and Young-Scholten, 2002). There is evidence that orthography can help to acquire new phonemic categories of the L2 (Escudero, Hayes-Harb, & Mitterer, 2008; Escudero, Simon, & Mulak, 2014; Showalter & Hayes-Harb, 2013; Simon, Chambless, & Kickhöfel Alves, 2010), but can also have a negative impact on the acquisition of L2 phonology (Bassetti, 2007; Bassetti & Atkinson, 2015; Young-Scholten & Langer, 2015), especially when the grapheme-phoneme relations of the L2 are different from the L1 (Escudero & Wanrooij, 2010; Hayes-Harb, Nicol, & Barker, 2010 et al., 2010). Furthermore, there is evidence that late L2 learners in instructed learning

environments can have an orthographic bias in their lexical knowledge, especially in the recognition of decontextualized word forms (Veivo, Suomela-Salmi, & Järvikivi, 2015). At the same time, words for these learners can have imprecise phonological representations (Cook & Gor, 2015; Cook, Pandža, Lancaster, & Gor, 2016), which may lead to the activation of false semantic content (Cook et al., 2016), but also to increased lexical competition (Broersma & Cutler, 2011). If the phonological representations of L2 words are more imprecise and unstable than those for L1 words, they may also be less well connected to their orthographic counterparts. As proficiency in the L2 increases, phonological representations are likely to become more accurate (Darcy, Daidone, & Kojima, 2013) and the orthographic bias in accessing semantic content decreases (Veivo et al., 2015). Taken together, the lexicon of late L2 learners in instructed learning could be orthographically biased so that orthographic representations may be more robust than phonological representations. Moreover, this relative bias might decrease as proficiency increases. In the present study, we evaluated this orthographic bias hypothesis by examining the flow of information from spoken word forms to written word forms in late L2 learners at different proficiency levels.

Previous studies have shown that proficiency can influence orthographic activation in L2 spoken word processing: Orthographic information during spoken word processing is activated more rapidly and more strongly in more proficient than in less proficient L2 learners (Mitsugi, 2016; Veivo & Järvikivi, 2013; Veivo, Järvikivi, Porretta, &

Hyönä, 2016). Specifically, Veivo et al. (2016) used the visual world paradigm with printed referents and observed a significant effect for the degree of orthographic overlap of the vowels in targets and competitors (target: *<mince>* /mɛ̃s/ 'slim' vs. O+ competitor: *<mite>* /mit/ 'moth' or O- competitor: *<mythe>* /mit/ 'mythe'), but only for higher proficiency participants. This suggests that orthographic information modulates L2 spoken word identification at least for higher proficiency learners. However, Veivo et al. (2016) did not contrast the two types of within-language L2 competitors in the same experiment or investigate the activation of between-language competitors from the participants' L1 to evaluate the activation of L1 orthography in L2 spoken word processing. The present study was designed to fill this gap.

Current study

In the present study, our main objectives were to investigate the mapping of spoken L2 words onto their written referents, and to evaluate whether this mapping is mediated mainly via orthographic or phonological information. For this purpose, we used the visual world eye-tracking paradigm (Allopenna, Magnuson, & Tanenhaus, 1998; Cooper, 1974; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; for a review, see Huettig, Rommers, & Meyer, 2011) in a task where spoken words are matched with their written counterparts (cf. McQueen & Viebahn, 2007; Huettig & McQueen, 2007; Huettig & McQueen, 2011). We studied Finnish learners of French with a wide range of proficiency levels. The task in both experiments consisted of listening to spoken instructions in

French ('*cliquez sur le mot cidre*') and clicking on one of the four words (target, competitor and two unrelated distractors) that appeared on the computer screen 200ms before the acoustic onset of the target word. The spoken target words were accompanied by a high orthographic low phonological overlap (OH-PL) competitor (e.g. *<cidre>* /sidʀ/ 'cider' vs. *<cintre>* /sɛ̃tʀ/ 'coat hanger') or a low orthographic high phonological overlap (OL-PH) competitor (e.g. *<cidre>* /sidʀ/ vs. *<cycle>* /sikl/ 'cycle') either in the L2 (Exp. 1) or in the L1 (Exp. 2).

If orthographic input in L2 acquisition leads to an orthographic bias in the lexical knowledge of late L2 learners (e.g. Young-Scholten, 2002; for a review, see Bassetti, 2008), we expect orthographically similar competitors to delay the mapping more than phonologically similar competitors. If the precision of phonological representations depends on proficiency (Darcy et al., 2013), proficiency will affect the speed of the mapping process. Based on previous results (Veivo & Järvikivi, 2013), lower proficiency learners might activate sublexical grapheme-phoneme correspondences of the L1, which would show as increased activation of phonologically similar L1 competitors in Exp. 2.

We started by investigating in Exp. 1 the matching of French spoken and written L2 word forms in the presence of within-language orthographic and phonological competitors.

Experiment 1

Method

Participants

Sixty-four students from the University of Turku participated for course credit or volunteered. They reported no hearing impairment or language deficits and had normal or corrected-to-normal vision. All participants were native speakers of Finnish who had learned French as a foreign language in instructed learning. None of the participants had acquired French or any other language besides Finnish before the age of three. Their age of onset for L2 French varied between 5 and 45 (median = 14). This means that they were all either literate or had started to acquire literacy in their L1 when they began to learn the L2. The participants represented a wide range of proficiency levels ranging from beginners to highly proficient. They evaluated their proficiency in French for five subskills (listening, reading, spoken interaction, spoken production, writing) with CEFR self-assessment grid (2001, p. 26-27). Each subskill was self-assessed on six levels which were assigned values from 1 to 6. The maximum score for proficiency for each participant was therefore 30[1]. Participant-related background information is summarized in Table 1.[2]

<Please, insert Table 1 about here>

Materials

The visual displays comprised four words: target, competitor and two distractors. There were 20 target words (e.g., *<cidre>*) each associated

[1] A comparison of self-ratings and DIALANG test scores (for DIALANG, see Huhta, Luoma, Oscarson, Sajavaara, Takala & Teasdale, 2002) in a previous study (Veivo, Suomela-Salmi & Järvikivi, 2015) showed that self-reported proficiency scores based on the CEFR-scale correlated highly with DIALANG test results.

[2] An L1 speaker control group (*n* = 24) completed Exp. 1 for control purposes. A linear mixed effect regression analysis of the looks to targets showed that the L1 group was marginally faster than the L2 group in finding the targets, and that the type of overlap in the L1 group was not significant.

with either a high orthographic low phonological (OH-PL) overlap competitor (e.g., *<cintre>*) or a low orthographic high phonological (OL-PH) overlap competitor (e.g., *<cycle>*). In the OH-PL condition, targets and competitors had a word initial orthographic overlap of two letters so that the nucleus vowel of the first syllable was always spelled similarly but pronounced differently (e.g. *<cidre>* /sidʀ/ 'cider' vs. *<cintre>* /sɛ̃tʀ/ 'coat hanger')[3]. In the OL-PH condition, targets and competitors always had a word initial phonological overlap of two sounds so that the nucleus vowel of the first syllable was pronounced similarly but spelled differently (e.g. *<cidre>* /sidʀ/ vs. *<cycle>* /sikl/ 'cycle'). Each target (e.g. *<cidre>* /sidʀ/) and its competitors (vs. *<cintre>* /sɛ̃tʀ/ and *<cycle>* /sikl/) were associated with two distractor words that were orthographically, phonologically and semantically unrelated. The two competitors were matched for frequency (Lexique 3; New, Pallier, Ferrand, & Matos, 2001) as well as possible. The mean frequency of the OH-PL –competitors was 43.7 per million and of the OL-PH –competitors 47.9 per million. Also, distractors in each display were matched for frequency with the target, 32.6 and 35.3 per million, respectively. Targets, competitors and distractors were also matched for written length[4]. The 20 target word sets are listed in Appendix 1. In addition to the target displays, 50 filler displays were constructed. In order to avoid the participants developing test-taking strategies and recognizing

---

[3] This was done because it was not possible to find competitors within the selection criteria with longer orthographic overlap. This is due to the relatively consistent grapheme to phoneme relations of French.
[4] To meet the selection criteria, a one letter difference in length between the words in the display was allowed.

the target displays on the basis of formal similarity between the words, 20 of the filler displays had an overlap between the distractor words. In 10 of these filler displays, the distractors had an OH-PL overlap, and in 10 displays, the distractors had an OL-PH overlap. The remaining 30 filler sets comprised four words with no orthographic, phonological or semantic overlap. In sum, Exp. 1 consisted of 70 trials (20 target word displays, 20 manipulated filler displays and 30 filler displays).

Each target word was embedded in a French sentence instructing the participant to click on the target word (e.g. *"cliquez sur le mot cidre"*). These sentences were recorded digitally using the SANAKO Lab100 hardware in the Learning, Age and Bilingualism laboratory (LAB-lab) at the University of Turku. A female native speaker of French, unaware of the aims of the study, read the sentences in a randomized order with a brief prosodic break before each target word. The mean duration for target words was 616ms.

Design and procedure

Each trial consisted of responding to the spoken instruction sentence, e.g. *"cliquez sur le mot cidre"*, by choosing the target word with a mouse click among the four words appearing on the computer screen. The position of each type of word was randomized for each display. For the target word displays, the competitors in the two experimental conditions were counterbalanced between two lists so that each list contained an equal number of OH-PL (10) and OL-PH (10) overlap competitors. The order for the presentation of the 70 trials was randomized

for each participant, and the participants were assigned to the two experimental lists in the order of appearance.

Participants' eye movements were monitored using a head-mounted SR EyeLink II eye-tracker (www.sr-research.com) sampling at 500Hz. Each trial started with drift correction where the participants fixated on a small cross appearing in the center of the screen for the experimenter to accept the gaze accuracy. After that, the spoken instruction to click on the target word was given via headphones. The visual display (see Figure 1) appeared on the screen 200ms before the onset of the target word (cf. McQueen & Viebahn, 2007; Huettig & McQueen, 2007; and Salverda & Tanenhaus, 2010). As it takes about 200ms to program and launch a saccade after a stimulus is presented (Matin, Shao, & Boff, 1993), this assured that the participants were not able to read the target words and have access to the phonological form via orthography before hearing the targets. The written words were presented in lowercase Times New Roman font being approximately 3 to 4° wide, with the center of each word appearing approximately 8° from the center of the screen (Figure 1).

<Please, insert Figure 1 about here>

Before the main experiment, participants were familiarized with the task by presenting a practice block of ten displays consisting of unrelated words. After that, they were presented with experiments 1 and 2. The order of the experiments was counterbalanced between participants.

Results and discussion

Five trials (0.5% of the data) were removed from the analyses

because the participants clicked on the competitor word instead of the target word.[5] The proportion of looks to the targets, to the competitors and to the distractors was determined for each trial and for each participant by calculating the number of looks to each word in 20ms time bins. Mean proportions of looks to each type of word in the two experimental conditions for a 1200ms period starting from target word onset are presented in Figure 2.

<Please, insert Figure 2 about here>

Proportions of looks to each type of word do not differ at word onset, but as Figure 2 shows, looks to distractors start to diverge from target and competitor looks in both experimental conditions at about 300ms after the onset of the target word. Looks to competitors increase until around 500ms, and looks to targets increase until reaching the asymptote around 1000ms post onset. Therefore, we examined the data more in detail within a time window ending at this latter time-point (200ms – 1000ms after target word onset). The proportions of fixations were logit-transformed for statistical analyses[6] (Fox & Weisberg, 2011), providing an unbounded measure in which zero represents 50% of looks (Barr, 2008).

Visual world eye-tracking data is inherently time-series data and usually presents nonlinearly over time (see Figure 2). Additionally, it is possible that the time-course interacts with other continuous variables, such as proficiency (cf. Veivo et al., 2016), which may also be nonlinear.

---

[5] Additionally, three targets were removed because of an error in choosing the OL-PH competitors.
[6] Logit-transformation applies a small adjustment to the values 0 and 1 (using 0.025 and 0.975 respectively).

We therefore used generalized additive mixed modelling (GAMM) (Baayen, Vasisth, Kliegel, & Bates, 2017; Hastie & Tibshirani, 1990; Wood, 2006), which does not assume a linear relationship between predictors and the response variable and is capable of modeling interactions between continuous variables (here, Time and Proficiency) (see also Baayen, Vasisth, Bates & Kliegl, 2015; Baayen, van Rij, de Cat & Wood, 2016; Veivo et al., 2016). Additionally, given the time series nature of the data, *GAMM* also allows for the control of autocorrelation in the data (see for instance Porretta, Kyröläinen, van Rij, & Järvikivi, 2018). Autocorrelation relates to the correlation between data points in a time series; a measurement at timepoint *t* is correlated to differing degrees with a measurement at timepoint *t-i*, depending on the lag. Autocorrelation is particularly problematic because it can greatly increase overconfidence of the model estimates.

In order to understand how online target word processing is modulated by proficiency and overlap, we modeled logit transformed Looks to the target word as a function of Time (200-1000ms after target onset), Proficiency (ranging from A1 to C2), and Overlap condition (OH-PL vs. OL-PH). Additionally, List and Trial were included in the analysis as control variables. Lastly, to control for individual variation in looking behavior, we created the variable Event. Here, Event represents the combination of Participant and Trial, capturing participants' variable responses to different items in the experiment. Event was included in the model as a random effect, allowing each unique time series to have its own

intercept in the model (Baayen et al., 2016; Nixon, van Rij, Mok, Baayen, & Chen, 2016; Porretta, Tucker, & Järvikivi, 2016).

It is reasonable to expect that Proficiency (a continuous variable) may influence the time-course of processing nonlinearly. To allow for this, we used a tensor product (Wood, 2006) for a non-linear relationship between Time and Proficiency. Further, also using a tensor product, a difference surface (Baayen, 2010; Wood, 2006) was included for Overlap condition. This approach allows for the evaluation of the significance of the factor relative to the interaction of Time and Proficiency. In this case, the difference surface informs how and where OH-PL is different from the overall effect by adding an additional smoothing parameter on top of the main trend (Zuur et al. 2009). Lastly, Trial order was included as a smooth term and List was included as a parametric term.

The model was fitted to the data through a series of steps in order to assess the contribution of each variable. First, we fitted a full model, i.e., all the predictors, as described above. Second, autocorrelation was estimated from the data (rho = 0.895, indicating a fairly high correlation between subsequent time points), and the model was refitted including this parameter to adjust the confidence of the estimates. Third, we evaluated the contribution of the individual predictors in the model. For this, two criteria were used: the *p*-value of the term (indicating whether a given effect is not zero) and Maximum Likelihood (ML) score comparison between model variants (indicating whether the inclusion of the predictor improved the fit of the model) (Zuur et al., 2009). This process was done

iteratively in a backward step-wise fashion until the model contained only predictors that were statistically significant and contributed to the model fit. Trial and the difference surface for Overlap Condition were removed through the fitting process, indicating that the order of presentation of the targets was not significant ($\chi^2(2) = 1.017$, p = 0.362), nor was the type of overlap between targets and competitors ($\chi^2(5) = 1.182$, p = 0.797).

ML score comparisons with Chi-square tests between variant models justified including Proficiency as an input variable ($\chi^2(3) = 30.192$, p < 0.001). The resulting model contained the following predictors: Event, Experimental List, Time * Proficiency, and explained 30.6% of the deviance. The statistics for the parametric and smooth terms of the model with the best fit are summarized in Table 2. The significant effect of Proficiency over Time is depicted in Figure 3.

<Please, insert Table 2 about here>

<Please, insert Figure 3 about here>

In interpreting the *GAMM* results, visual inspection of the figures is essential, perhaps even more so than in other types of data analysis. Figure 3 presents the interaction between Proficiency and Time as a regression surface, showing that overall, as time progressed, participants were generally more likely to look at the target. Here darker shades of grey represent less looks to the target, whereas lighter shades of grey represent more looks to the target, and the contour lines indicate the rate of change.

More interestingly, as proficiency increased, the participants were more likely to look at the target. Importantly, lower proficiency learners

looked at the targets later than higher proficiency learners. Proficiency especially influenced processing in participants with proficiency scores under 15 (equal to CEFR-levels A1, A2 and B1) and did so in a graded fashion. For example, if we follow the time-course for participants with proficiency scores 5 and 20, we find that lower proficiency participants were less likely to fixate the targets (between 400ms and 600ms). However, we can also see that the effect of proficiency was not linear along the proficiency continuum. This is evidenced by the shape of the contour lines which indicate a strong effect of proficiency for participants with scores under 15, and little-to-no effect for participants with scores over 15.

The results of Exp. 1 fail to provide evidence that the OH-PL overlap between targets and within-language competitors delays the mapping between spoken and written forms more than OL-PH overlap. This suggests that when both orthographic and phonological competitors are present at the same time, both orthographic and phonological information is used in the matching process to a similar degree. We will return to this issue in detail in the general discussion. However, our results confirm that the speed of target identification depends on L2 proficiency in a non-linear fashion: more proficient L2 listeners fixate the targets faster than less proficient learners, and the influence of proficiency is more pronounced in the lower half of the proficiency scale.

As we were interested in how the L1 modulates L2 performance, we next moved on to investigate the activation of orthographic and

phonological information from the participants' L1 in the recognition of L2 word forms in Exp. 2. This experiment was designed to examine the impact of L1 orthography on the mapping process of the L2 at different proficiency levels.

Experiment 2

Method

Participants

Participants were the same as in Exp. 1.[7] Therefore the two experiments were presented in a counterbalanced order between the participants.

Materials

As in Exp. 1, the visual displays consisted of four words: the target, the competitor and two distractor words. Exp. 2 comprised 28 target words. Each French target word (e.g., *<paume>* /pom/*,* 'palm') was associated with a Finnish competitor with either a high orthographic low phonological (OH-PL) overlap (e.g., *<pauhu>* /pauhu/ 'roar') or low orthographic high phonological (OL-PH) overlap competitor (e.g., *<pommi>* /pom:i/ 'bomb'). In the OH-PL condition, targets and competitors had a word initial orthographic overlap of three letters so that the nucleus vowel of the first syllable was always spelled similarly but pronounced differently according to L1 phoneme-grapheme correspondences (cf. *<paume>* /pom/

---

[7] The same L1 control group who participated in Exp. 1, also participated in Exp. 2 for continuity, but the data were not analysed in more detail because the participants did not have any knowledge of Finnish, the language of the competitor words.

vs. *<pauhu>* /pauhu/[8]. In the OL-PH condition targets and competitors always had a word initial phonological overlap of two sounds so that the nucleus vowel of the first syllable was always pronounced as similarly as possible (the vowel systems of the two languages not being the same) but spelled differently (cf. *<paume>* /pom/ vs. *<pommi>* /pom:i/). Each target (e.g., *<paume>* /pom/) and its competitors (e.g. *<pauhu>* /pauhu/ and *<pommi>* /pom:i/) were associated with two orthographically, phonologically and semantically unrelated distractors, one in the L1 and the other in the L2. The visual display therefore always consisted of two words in the L2 (French) - the target and the first distractor - and of two words in the L1 (Finnish) - the competitor and the second distractor. Competitors were matched as well as possible for written frequencies reported in the unpublished *Turun Sanomat* newspaper lexical database (comprising 22.7 million word tokens) using the WordMill Lexical Search Program (Laine & Virtanen, 1999). The mean frequency was 12.6 per million for the OH-PL overlap competitors and 29.1 per million for the OL-PH overlap condition. Distractors were matched for frequency with the target. The mean frequency of the target words was 118.8 per million; the mean frequency of the distractor words was 109.7 per million (101.8 per million for the L1 distractors and 117.6 per million for the L2 distractors). The 28 target word sets are listed in Appendix 2. In addition to the target set, we constructed a 28 item filler set that consisted of four words with no

---

[8] It was not possible to find competitors with equal syllable length in Finnish. It is therefore possible that if the participants knew any one-syllable target words, it may have helped to overrule competitors with two syllables.

orthographic, phonological or semantic overlap. In the filler sets, the target and another word were in the L2 (French) and two additional words in the L1 (Finnish). The procedure for recording the targets was as in Exp. 1. The mean duration of target words was 592ms.

Design and Procedure

The design and procedure were identical to those in Exp. 1.

Results and discussion

Before the analyses, 6 trials (0.3% of the data) were removed from the data because the participants erroneously clicked on the competitor word. As in Exp. 1, the proportion of looks to the targets, to the competitors and to the distractors was determined for each trial and for each participant by calculating the number of looks to each type of word in 20ms time bins. The mean proportion of looks to each word in the display for 1200 ms, starting from the target word onset, is depicted in Figure 4.

<Please, insert Figure 4 about here>

Visual inspection of the plots shows that looks to distractors start to diverge from competitor and target looks at about 300ms and that looks to competitors start to decline around 600ms post target word onset. Looks to targets increase until reaching the asymptote around 1000ms. As in Exp. 1, the proportions of looks in each 20ms bin were logit-transformed for statistical analyses to give an unbounded measure.

As in Exp. 1, we examined the time course of target identification more in detail in a time window from 200ms until 1000ms after the target word onset. Again, we used *GAMM* and the model was structured exactly

as in the analysis of Exp. 1. The model was fitted using the same steps and procedure as in Exp. 1. Through this process, an autocorrelation parameter of rho = 0.895 was included and Trial was removed from the input variables. ML score comparisons with Chi-square tests between variant models supported the inclusion of List ($\chi^2(1)$ = 6.599, p < 0.001), the difference surface for Overlap Condition ($\chi^2(6)$ = 11.860, p < 0.001) and Proficiency ($\chi^2(6)$ = 27.278, p < 0.001) as input variables. The resulting model consisted of random intercepts for Event, a parametric term for List, an interaction between Time and Proficiency, and a difference surface for Overlap Condition. This final model explains 38% of the deviance; the estimates for the parametric and smooth terms are summarized in Table 3.

<Please, insert Table 3 about here>

<Please, insert Figure 5 about here>

Panel 1 of Figure 5 presents the interaction of Proficiency and Time as a regression surface. As in the results of Exp. 1, as time progressed, participants were generally more likely to look at the target. Again, lighter grey represent greater likelihood of target looks, while darker grey represents lesser likelihood of target looks, and the contour lines indicate the rate of change. Similar to Exp. 1, as proficiency increases, the participants were more likely to look at the target; lower proficiency learners were less likely to look at the target over time than higher proficiency learners. However, if we compare the shape of the contour lines in figures 3 and 5 (Panel 1), we can see that this effect in Exp. 2 is less pronounced than in Exp. 1, affecting primarily the lowest proficiency

learners with proficiency scores from 5 to 10 (CEFR levels A1 and A2).

In contrast to Exp. 1, we also observed a significant adjustment due to Overlap Condition (i.e., significant difference surface), indicating that the Time*Proficiency surface for OH-PL (Figure 5, Panel 2) deviates significantly from the main Time*Proficiency surface (Figure 5, Panel 1). Between the two panels we see a greater influence of proficiency in the second half of the time-course. As explained above, Panel 1 of Figure 5 presents the effect of Proficiency over Time for Target Looks in the whole data. To aid in the visualization of the effect of Overlap condition, Panel 2 depicts the same effect but with the difference surface added to show the effect in the OH-PL-condition (e.g. *<paume>* /pom/ vs. *<pauhu>* /pauhu/). Compared to Panel 1, Panel 2 shows that, in the second half of the time-course, participants with a proficiency score above 15 (CEFR-levels B2, C1 and C2) were more likely to fixate the target, whereas participants with proficiency scores below 15 (CEFR-levels A1, A2 and B1) were less likely to fixate the target. The significant difference surface for Overlap ($F(8.425,72249.49) = 3.839$, $p < 0.001$) indicated that this adjustment across proficiency was not zero after approximately 600ms post target word onset.

Thus, when presented with orthographic competitors from the participants' L1, higher proficiency learners were more likely to look at the target while lower proficiency learners were less likely to look at the target. Also, learners in the upper half of the proficiency scale behaved more uniformly than learners in the lower half of the proficiency scale. In

other words, lower proficiency learners showed more variation in the speed of finding the targets across proficiency scores.

The results of Exp. 2 indicate that the mapping between L2 spoken and written word forms is influenced by between-language competitors from the participants' L1. This finding is in line with previous findings on language non-selective lexical access in L2 auditory word recognition (e.g. Lagrou, Hartsuiker, & Duyck, 2011). As in Exp. 1, the effect of proficiency on matching spoken and written word forms was more pronounced in lower proficiency participants. Furthermore, analysis of the time-course of activation in the matching process with *GAMM* revealed an effect of L1 orthographic overlap that appeared 600ms after target word onset: compared to the main trend, it was more difficult for the participants in the lower half of the proficiency scale to find the targets (e.g. *<paume>* /pom/) in the presence of OH-PL L1 competitors (e.g. *<pauhu>* /pauhu/). The effect of proficiency was also more pronounced in the presence of OH-PL L1 competitors for lower proficiency participants in this same time-window.

These results suggest that in spoken word identification in late L2 learners, both orthographically and phonologically similar L1 words are activated early in the recognition process, but later on, orthographic information is activated more than phonological information. In contrast to Veivo and Järvikivi, (2013), we did not find evidence for the effect of L1 orthography depending on proficiency. It is likely that, unlike in masked priming that taps into the early phase of processing, in the kind of

matching task used in the present study, lower proficiency learners did not rely on sublexical correspondences of the L1 more than on those of the L2. Instead, we observed significantly more activation for orthographically similar L1 words at all proficiency levels. Taken together, these results speak for an orthographically mediated activation in matching the spoken and written word forms, compatible with the orthographic bias hypothesis presented above.

General discussion

Our objective in the present study was to investigate the impact of orthographic and phonological information in the matching of spoken and written L2 word forms. This was done in the presence of within-language (Exp. 1) and between-language competitors (Exp. 2). We studied L1 Finnish late learners of L2 French and used the visual world eye-tracking paradigm with printed words as referents. We set out to evaluate the orthographic bias hypothesis, according to which orthographic representations in late L2 learners are more robust than phonological representations. This hypothesis makes the prediction that spoken L2 target words activate mainly orthographic representations, and therefore orthographically similar competitors in the L2 and L1 should delay the mapping between spoken and written L2 word forms more than phonologically similar competitors. Because phonological representations are likely to become more accurate as proficiency increases, proficiency was expected to affect the speed of target identification. For L1 competitors, we predicted that higher proficiency learners might show

orthographic effects, whereas lower proficiency learners might rely on L1 sublexical correspondences, which would show as increased activation of phonologically similar L1 competitors.

First of all, the results of our within-language experiment (Exp. 1) showed that orthographically and phonologically similar L2 word forms are activated rapidly, around 300ms after target word onset. This result is in line with previous visual world studies reporting phonological effects (e.g. Huettig & McQueen, 2007) and orthographic effects in L1 speakers (Salverda & Tanenhaus, 2010). The results of our between-language experiment (Exp. 2) showed that spoken L2 words also activated orthographically and phonologically similar word forms in the participants' L1, Finnish.

In the within-language experiment (Exp. 1), competitors with an OH-PL overlap did not delay target identification more than competitors with an OL-PH overlap. Written competitors in the OH-PL condition were selected so that the vowel of the nucleus was always pronounced differently from the targets (/sidʀ/ <cidre> vs. /sɛ̃tʀ/ <cintre>), whereas written competitors in the OL-PH condition were always selected so that the vowel of the nucleus was pronounced similarly to the targets (/sidʀ/ <cidre> vs. /sikl/ <cycle>). If the participants had unstable grapheme-phoneme correspondences in the L2, the competitors in the two overlap conditions were not necessarily phonologically different for them. Also, because the two competitors were allowed to share word-final letters (like <e>), the competitors may not have been orthographically different enough

to produce a significant difference. This result does not necessarily rule out the orthographic bias hypothesis. It is possible that the activation spreads mainly via orthographic representations. In other words, a spoken word like /sidʀ/ activates its written counterpart *<cidre>*, which then sends activation to both *<cintre>* and *<cycle>*. This is likely especially if the learners are not aware of the pronunciation difference between the two written competitors.

In the between-language experiment (Exp. 2), the analysis revealed a non-linear interaction between time, proficiency, and the type of overlap. When spoken L2 targets like /pom/ *<paume>* were presented with OH-PL competitors from the participants' L1 (*<pauhu>* / pauhu/), these competitors delayed target identification significantly compared to the main trend in a time window from 600ms to 1000ms post-target onset for lower proficiency participants. In addition, the effect of proficiency was more prominent in presence of these OH-PL competitors for these participants in the same time window. Unlike in Exp. 1, the degree of phonological overlap between targets and competitors in Exp. 2 was unambiguous for participants at all proficiency levels, because the competitors were in the L1. According to L1 pronunciation rules, there was always a phonological mismatch between L2 targets and L1 OH-PL competitors. Therefore, spoken L2 words like /pom/ *<paume>* could activate OH-PL competitors like /pauhu/ *<pauhu>* only via orthographic

representations[9]. As predicted by the orthographic bias hypothesis, orthographic representations are activated significantly more than phonological representations at this relatively late time point.

Unlike we assumed on the basis of previous results (Veivo & Järvikivi, 2013), lower proficiency learners were not relying on L1 grapheme-phoneme correspondences in Exp. 2, because written L1 competitors which could be pronounced similarly to the targets did not delay the matching. Instead, the delay in target word recognition due to orthographically similar L1 competitors was even more salient for lower proficiency learners. This is in line with the assumption that the orthographic bias would be more prominent in lower proficiency learners and decrease as proficiency increases. It is also possible that the higher proficiency learners were more able to suppress the irrelevant between-language information and that this suppression was enhanced when the orthographic information made it clearer what the unfolding word form would be. This interpretation is in line with the results of Blumenfeld and Marian (2013) which suggest that more efficient cognitive control is associated with reduced cross-linguistic activation in bilinguals in a relatively late phase of lexical competition (633-767ms).

Based on previous research (e.g. Mitsugi, 2016; Veivo & Järvikivi, 2013; Veivo et al., 2016), we assumed that proficiency would influence the

---

[9] Because grapheme to phoneme correspondences in Finnish are different from French, there was more variation in the OL-PH competitors in the between-language experiment (Exp. 2) than in the within-language experiment (Exp. 1). Whether this may impact participants to rely more on orthography should be studied in a subsequent experiment contrasting OL-PH competitors from both languages.

speed of the matching process. Our results confirmed this assumption: in both experiments, higher proficiency learners identified the targets significantly faster than lower proficiency learners. This difference in processing speed suggests that lower proficiency learners have less precise phonological representations to base the matching on (cf. Cook et al. 2016). Interestingly, the effect of proficiency on the speed of looking at the targets was not linear throughout the proficiency continuum: not only were the lower proficiency learners generally slower in looking at the target than higher proficiency learners, but they were also relatively slower the less proficient they were, unlike learners in the upper half of the proficiency continuum (CEFR-levels B2, C1 and C2) who behaved more homogenously in the task. This suggests that until this point, connections between the modalities develop gradually, but that above this point, they have mostly been acquired.

The results of the present study are discussed above in terms of co-activation orthographic and phonological information in spoken word processing (see e.g. Grainger et al., 2003). This same principle underlies the structure of the Bilingual Language Interaction Network for Comprehension of Speech (BLINCS) (Shook & Marian, 2013). In BLINCS, different levels of representation are represented by self-organizing maps which structure language input from both languages according to similarity between units. Further, ortholexical and phonolexical levels of representations are shared between both languages and interconnected via bi-directional links. In this respect, the architecture

of BLINCS is similar to the interactive activation models of bilingual written word recognition like BIA+ (Dijkstra & van Heuven, 2002).

To our knowledge, there are no current models of L2 spoken word recognition which would comprise a possibility for interaction between the modalities. The BLINCS model was designed to model spoken word recognition in bilingual speakers, but it could be used to describe the co-activation of orthographic information in spoken word recognition also in late L2 learners. The results of the current study suggest that there are two important proficiency related features which should be considered in modeling L2 spoken word recognition: the relative strength of ortholexical and phonolexical representations as well as the degree of activation of L1 orthography are different depending on the level of L2 proficiency. Although BLINCS assumes that both languages are known equally well, its architecture based on the self-organizing principle fits well for describing also L2 spoken word recognition. Lower proficiency learners have less experience with L2 word forms, especially in speaking, and this lower frequency in the input causes the phonolexical representations to become weaker, or fuzzier (see Cook et al., 2016), than ortholexical representations.

The results of our Exp. 2 indicate that lower proficiency learners are also less efficient in inhibiting the activation of L1 word forms. The results of Veivo and Järvikivi (2013) suggest that this effect may rise from the sublexical phoneme level. Further, this finding suggests that L2 spoken language processing mechanisms could be qualitatively different at

different proficiency levels. In sum, our results show that orthography and L2 proficiency are factors that should be incorporated in any model of L2 spoken word recognition.

In the present study, we compared the use of orthographic and phonological information in the mapping of spoken and written L2 word forms. Based on studies on the importance of orthographic input in L2 acquisition (e.g. Young-Scholten, 2002), it was hypothesized that there is an orthographic bias in late L2 learners' lexicons with orthographic representations being more robust and accurate than phonological representations. This bias predicts increased orthographic activation in L2 spoken word processing compared to phonological activation. The results of our between-language experiment are compatible with this hypothesis because orthographically similar L1 words delayed the matching process significantly more than phonologically similar L1 words. However, the results of our within-language experiment did not provide straightforward evidence in support of the orthographic bias, and therefore it should be evaluated in subsequent research. Finally, our results suggest that the competences needed for combining phonological and orthographic word forms develop significantly more in the lower half than in the upper half of the proficiency scale. Whether the findings on activation of orthographic information apply to more naturalistic L2 spoken word processing contexts remains a question for future research.

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38,* 419–439.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge, UK: Cambridge University Press.

Baayen, R. H. (2010). The directed compound graph of English. An exploration of lexical connectivity and its processing consequences. In S. Olson (Ed.), *New impulses in word-formation (Linguistische Berichte Sonderheft 17),* pp. 383–402. Hamburg: Buske.

Baayen, R. H., van Rij, J., de Cat, C., & Wood, S. N. (2016). Autocorrelated errors in experimental data in the language sciences: Some solutions offered by generalized additive mixed models. In. D. Speelman, K. Heylan, & D. Geeraerts (Eds.), *Mixed effects regression models in linguistics*. Berlin: Springer. Forthcoming. arXiv preprint arXiv:1601.02043.

Baayen, R. H., Vasishth, S., Kliegl, R., & Bates, D. (2017). The cave of shadows: Addressing the human factor with generalized additive mixed models. *Journal of Memory and Language*, 94, 206–234.

Barr, D. J. (2008). Pragmatic expectations and linguistic evidence: Listeners anticipate but do not integrate common ground. *Cognition, 109*, 18–40.

Bartolotti, J., Daniel, N. L., & Marian, V. (2013). Spoken words activate cross-linguistic orthographic competitors in the absence of phonological overlap. *Proceedings of the 35th Annual Conference of the Cognitive Science Society*, pp. 1828–1832. Austin TX: Cognitive Science Society.

Bassetti, B. (2006). Orthographic Input and Phonological Representations in Learners of Chinese as a Foreign Language. *Written Language and Literacy, 9,* 95–114.

Bassetti, B. (2007). Effects of hanyu pinyin on pronunciation in learners of Chinese as a foreign language. In A. Guder, X. Jiang & Y. Wan, (Eds.), *The cognition, learning and teaching of chinese characters*. Beijing, China: Beijing Language and Culture University Press.

Bassetti, B. (2008). Orthographic input and second language phonology. In P. Thorsten and M. Young-Scholten (Eds.), *Input matters in SLA*. pp. 191-206. Clevedon, UK, Multilingual Matters.

Bassetti, B., & Atkinson, N. (2015). Effects of orthographic forms on pronunciation in experienced instructed second language learners. *Applied Psycholinguistics, 36*(1), 67-91.

Best, C., & Tyler, M. (2007). Nonnative and Second-Language Speech Perception, Commonalities and Complementarities. In O.-S. Bohn & M. Munro (Eds.), *Language Experience in Second Language*

Speech Learning: In Honor of James Emil Flege, pp. 13–34. Amsterdam: John Benjamins Publishing Company.

Blumenfeld, H. K., & Marian, V. (2007). Constraints on parallel activation in bilingual spoken language processing: Examining proficiency and lexical status using eyetracking. *Language and Cognitive Processes, 22*, 633–660.

Blumenfeld, H. K., & Marian, V. (2013). Parallel language activation and cognitive control during spoken word recognition in bilinguals. *Journal of Cognitive Psychology, 25*, 547–567.

Broersma, M., & Cutler, A. (2011). Competition dynamics of second-language listening. *Quarterly Journal of Experimental Psychology, 64,* 74–95.

CEFR, *Common European Framework of Reference for Languages: Learning, Teaching, Assessment.* (2001). Cambridge: Cambridge University Press.

Cook, S. V., & Gor, K. (2015). Lexical access in L2: Representational deficit or processing constraint? *The Mental Lexicon*, *10,* 247–270.

Cook, S. V., Pandža, N. B., Lancaster, A K., & Gor, K. (2016). Fuzzy nonnative phonolexical representations lead to fuzzy form-to-meaning mappings. *Frontiers in Psychology*, *7*, 1-17. DOI: 10.3389/fpsyg.2016.01345.

Cooper, R. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of

speech perception, memory, and language processing. *Cognitive Psychology, 6,* 84–107.

Darcy, I., Daidone, D., & Kojima, C. (2013). Asymmetric lexical access and fuzzy lexical representations in second language learners. *The Mental Lexicon, 8*, 372–420.

Dijkstra, T., & van Heuven, W. J .B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition*, 5, 175–197.

Escudero, P., & Wanrooij, K. (2010). The effect of L1 orthography on non-native vowel perception. *Language and Speech, 53*, 343–365.

Escudero, P., Simon, E., & Mulak, K. (2014). Learning words in a new language: Orthography doesn't always help. *Bilingualism: Language and Cognition, 17*(2), 384–395.

Escudero, P., Hayes-Harb, R., & Mitterer, H. (2008). Novel second-language words and asymmetric lexical access. *Journal of Phonetics, 36*, 345–360.

Fox, J. & Weisberg, S. (2011). *An {R} companion to applied regression*, Second Edition. Thousand Oaks CA: Sage.

Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin, 123*, 71–99.

Frost, R., & Ziegler, J. C. (2007). Speech and spelling interaction: The interdependence of visual and auditory word recognition. In M. G.

Gaskell (Ed.), *The Oxford handbook of psycholinguistics,* pp.
107–118. Oxford, England: Oxford University Press.

Grainger, J., Diependaele, K., Spinelli, E., Ferrand, L., & Farioli, F.
(2003). Masked repetition and phonological priming within and
across modalities. *Journal of Experimental Psychology: Learning,
Memory, and Cognition, 29*, 1256–1269.

Grainger, J., & Ferrand, L. (1996). Masked orthographic and phonological
priming in visual word recognition and naming: cross-task
comparisons. *Journal of Memory and Language, 35*, 623– 647.

Hastie, T. & Tibshirani, R. (1990). Generalized additive models.
Monographs on Statistics & Applied Probability, 43. Boca Raton:
Chapman and Hall/CRC.

Hayes-Harb, R., Nicol. J., & Barker, J. (2010). Learning the phonological
forms of new words: Effects of orthographic and auditory input.
Language & Speech, 53, 367–381.

Huettig, F., & McQueen, J. M. (2007). The tug of war between
phonological, semantic and shape information in language-
mediated visual search. *Journal of Memory and Language, 57*,
460–482.

Huettig, F., & McQueen, J. M. (2011). The nature of the visual
environment induces implicit biases during language-mediated
visual search. *Memory & Cognition*, *39*, 1068–1084.

Huettig, F., Rommers, J., & Meyer, A.S. (2011). Using the visual world paradigm to study language processing: a review and critical evaluation. *Acta Psychologica, 137*, 151–171.

Huhta, A., Luoma, S., Oscarson, M., Sajavaara, K., Takala, S. and Teasdale, A. (2002). DIALANG: a diagnostic language assessment system for learners. In J. C. Alderson (Ed.) *Common European Framework of Reference for Languages: learning, teaching, assessment. Case studies* (130–145). Strasbourg: Council of Europe.

Lagrou, E., Hartsuiker, R.J., & Duyck, W. (2011). Knowledge of a second language influences auditory word recognition in the native language. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 952–965.

Laine, M., & Virtanen, P. (1999). *WordMill lexical search program* [Computer software]. Turku, Finland: Centre for Cognitive Neuroscience, University of Turku.

Marian, V., & Spivey, M. (2003a). Competing activation in bilingual language processing: Within- and between language competition. *Bilingualism: Language and Cognition, 6*, 97–115.

Marian, V., & Spivey, M. (2003b). Bilingual and monolingual processing of competing lexical items. *Applied Psycholinguistics, 24*, 173–193.

Matin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead:

    Information-processing time with and without saccades.

    Perception & Psychophysics, *53*, 372–380.

Mitsugi, S. (2016). Proficiency influences orthographic activations during

    L2 spoken-word recognition. *International Journal of*

    Bilingualism. 1-16. DOI: 10.1177/1367006916666388

McQueen, J. M., & Viebahn, M. C. (2007). Tracking the recognition of

    spoken words by tracking looks to printed words. *Quarterly*

    Journal of Experimental Psychology, 60, 661–671.

New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données

    lexicales du français contemporain sur internet: LEXIQUE,

    L'Année Psychologique, 101, 447–462. http://www.lexique.org

Nixon, J. S., van Rij, J., Mok, P., Baayen, R. H., & Chen, Y. (2016). The

    temporal dynamics of perceptual uncertainty: eye movement

    evidence from Cantonese segment and tone perception. *Journal of*

    Memory and Language, 90, 103–125.

Porretta, V., Tucker, B., & Järvikivi, J. (2016). The influence of gradient

    foreign accentedness and listener experience on word recognition.

    Journal of Phonetics, 58, 1–21.

Porretta, V., A. Kyröläinen, J. van Rij and J. Järvikivi (2018). Visual world

    paradigm data: From preprocessing to nonlinear time-course

    analysis. In I. Czarnowski, R. J. Howlett & L. C. Jain. (Eds.)

    Intelligent Decision Technologies 2017. Proceedings of the 9th

    KES International Conference on Intelligent Decision

Technologies (KES-IDT 2017) - Part II. Cham: Springer

International Publishing, pp. 268–277. DOI: 10.1007/978-3-319-

59424-8_25.

Salverda, A. P., & Tanenhaus, M. K. (2010). Tracking the time course of

orthographic information in spoken-word recognition. *Journal of

Experimental Psychology: Learning, Memory, and Cognition, 5*,

1108–1117.

Seidenberg, M., & Tanenhaus, M. (1979). Orthographic effects in rhyme

monitoring, *Journal of Experimental Psychology: Human

Learning and Memory, 5*, 546–554.

Showalter. C. E., & Hayes-Harb, R. (2013). Unfamiliar orthographic

information and second language word learning: A novel lexicon

study. *Second Language Research*, 29, 185–200.

Showalter. C. E., & Hayes-Harb, R. (2015). Native English speakers

learning Arabic: The influence of novel orthographic information

on second language phonological acquisition. *Applied

Psycholinguistics, 36*, 23–42.

Shook, A., & Marian, V. (2013). The bilingual language interaction

network for comprehension of speech. *Bilingualism: Language

and Cognition, 16*, 304–324.

Simon, E., Chambless, D., & Kickhöfel Alves, U. (2010). Understanding

the role of orthography in the acquisition of a non-native vowel

contrast. *Language Sciences, 32*(3), 380–394.

Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science, 10*, 281–284.

Taft, M., Castles, A., Davis, C., Lazendic, G., & Nguyen-Hoan, M. (2008). Automatic activation of orthography in spoken word recognition: Pseudohomograph priming. *Journal of Memory and Language, 58*, 366–379.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*, 1632–1634.

Veivo, O., & Järvikivi, J. (2013). Proficiency modulates early orthographic and phonological processing in L2 spoken word recognition. *Bilingualism: Language and Cognition, 16*, 864–883.

Veivo, O., Suomela-Salmi, E., & Järvikivi, J. (2015). Orthographic bias in L3 lexical knowledge: Learner-related and lexical factors. *Language, Interaction and Acquisition, 6*, 271–294.

Veivo, O., Porretta, V., Järvikivi, J. & Hyönä, J. (2016). Orthographic activation in L2 spoken word recognition depends on proficiency: evidence from eye-tracking. *Frontiers in Psychology*. DOI: 10.3389/fpsyg.2016.01120.

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*, 57–95.

Wood, S. N. (2006). *Generalized additive models: An introduction with R*. Boca Raton: Chapman & Hall/CRC Press.

Young-Scholten M. (2002). Orthographic input in L2 phonological development. In P. Burmeister, T., Piske, & A. Rohde, (Eds.). *An Integrated View of Language Development: Papers in Honor of Henning Wode*. Trier: Wissenschaftlicher Verlag Trier, pp.263-279.

Young-Scholten, M., & Langer, M. (2015). The role of orthographic input in second language German: Evidence from naturalistic adult learners' production. *Applied Psycholinguistics, 36*(1), 93–114.

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

Zuur, A., Ieno, E., Walker, N., Saveliev, A., & Smith, G. (2009). *Mixed effects models and extensions in ecology with R*. New York: Springer.

Appendix 1: Targets and competitors in Exp. 1.

| | Target | | | OH-PL-competitor | | | OL-PH-competitor | |
ORT	PHON	freq.	ORT	PHON	freq.	ORT	PHON	freq.
cidre	sidʀ	3.99	cintre	sɛ̃tʀ	4.86	cycle	sikl	5.81
faute	fɔt	95.2	farce	faʀs	12.5	forge	fɔʀʒ	8.58
folie	foli	60.74	foire	fwaʀ	15.95	fauve	fov	13.4
lent	lã	64.12	leçon	ləsɔ̃	48.72	lampe	lãp	93.1
lynx	lɛ̃ks	1.22	lyre	liʀ	1.42	limbe	lɛ̃b	3.18
maure	mɔʀ	1.15	marle	maʀl	1.55	morve	mɔʀv	1.96
membre	mãbʀ	64.86	meuble	mœbl	53.99	manque	mãk	37.2
motte	mɔt	9.19	moule	mul	9.26	mauve	mov	17.6
impôt	ɛ̃po	6.55	imité	imite	1.82	ainsi	ɛ̃si	469
foie	fwa	17.97	fort	fɔʀ	212.8	foyer	fwaje	30.9
saule	sol	8.65	sapin	sapɛ̃	26.28	sonné	sone	3.58
sens	sãs	217.5	selon	səlɔ̃	110.9	sang	sã	207
serpe	sɛʀp	4.32	senti	sãti	1.76	saisi	sezi	3.92
signal	siɲal	23.11	simple	sɛ̃pl	148.6	syrien	siʀjɛ̃	6.82
singe	sɛ̃ʒ	22.57	sirop	siʀo	8.18	sympa	sɛ̃pa	7.77
solide	solid	42.77	source	suʀs	49.19	saumon	somɔ̃	4.73
vanter	vãte	23.24	vaste	vast	71.76	vendu	vãdy	2.3
mean		35.32			43.69			47.9

Appendix 2: Targets and competitors in Exp. 2.

Target			OH-PL-competitor			OL-PH-competitor		
ORT	PHON	freq.	ORT	PHON	freq.	ORT	PHON	freq.
hausse	os	2.36	haukka	hɑuk:ɑ	17.22	ostos	ostos	24.41
renom	ʀənɔ̃	2.43	renki	reŋki	4.89	rönsy	rønsy	3.79
roulis	ʀuli	3.04	rousku	rousku	3.79	ruska	ruskɑ	6.39
renvoi	ʀɑ̃vwa	4.8	rengas	reŋɑs	44.27	ranta	rantɑ	170.44
horde	ɔʀd	7.03	horros	hor:os	1.54	orja	orja	5.77
noix	nwa	12.23	noita	noitɑ	11.15	nuoli	nuoli	7.53
sauce	sos	13.72	sauma	sɑumɑ	9.96	solmu	solmu	13.88
pince	pɛ̃s	14.73	pinja	pinja	2.03	päästö	pæ:stø	40.18
moulin	mulɛ̃	19.05	moukka	mouk:ɑ	0.79	multa	multɑ	14.58
jouet	ʒwɛ	21.01	jousi	jousi	10.93	suure	su:re	24.63
souris	suʀi	22.57	soutu	soutu	6.87	surma	surmɑ	86.65
poupée	pupe	27.57	pouta	poutɑ	1.45	puppu	pup:u	2.38
paume	pom	35.47	pauhu	pɑuhu	1.98	pommi	pom:i	45.51
rumeur	ʀymœʀ	37.97	rumpu	rumpu	9.91	rytmi	rytmi	26.48
ruine	ʀɥin	39.39	ruiske	ruiske	1.41	ryijy	ryijy	3.61
voile	vwal	48.31	voide	voide	2.42	vuohi	vuohi	3.39
ennui	ɑ̃nɥi	56.62	enne	enne	7.00	anti	ɑnti	46.96
poil	pwal	76.01	poiju	poiju	2.11	puomi	puomi	5.46
neige	nɛʒ	80.88	neiti	neiti	5.90	neste	neste	66.65
toile	twal	106.62	toive	toive	90.13	tuomi	tuomi	36.70
peuple	pøpl	107.3	peukku	peuk:u	0.84	pöpö	pøpø	1.67
ventre	vɑ̃tʀ	141.96	ventti	vent:i	0.31	vanne	van:e	3.22
jardin	ʒaʀdɛ̃	185.81	jarru	jɑr:u	15.90	sarka	sarkɑ	13.66
route	ʀut	288.04	rouva	rouvɑ	54.19	rutto	rut:o	9.47
raison	ʀɛzɔ̃	308.78	raivo	rɑivo	11.59	ressu	res:u	3.79
peur	pøʀ	311.69	peura	peurɑ	7.36	pörssi	pørs:i	43.96
voix	vwa	612.7	voima	voimɑ	477.89	vuota	vuotɑ	7.53
nuit	nɥi	738.24	nuija	nuijɑ	7.49	nyöri	nyøri	2.11
		118.79			28.98			25.74