



RESEARCH REPORT 

# On the effects of task focus and processing level on the perception–production link in second-language speech learning

Miquel Llompart<sup>1,2</sup> 

<sup>1</sup>Department of Translation and Language Sciences, Universitat Pompeu Fabra, Barcelona, Spain; <sup>2</sup>Friedrich Alexander University Erlangen-Nuremberg, Erlangen, Germany

Email: miguel.llompart@upf.edu

(Received 01 December 2022; Revised 06 July 2023; Accepted 20 July 2023)

## Abstract

This study presents a reanalysis of existing data to investigate whether a relationship between perception and production abilities regarding a challenging second-language (L2) phonological contrast is observable (a) when both modalities must rely on accessing stored lexical representations and (b) when there is an asymmetry in task focus between perception and production. In the original studies, German learners of English were tested on their mastery of the English /ɛ/-/æ/ contrast in an auditory lexical decision task with phonological substitutions, a word-reading task, and a segmentally focused imitation task. Results showed that accurate nonword rejection in the lexical decision task was predicted by the Euclidean distance between the two vowels in word reading but not in imitation. These results extend previous findings to lexical perception and production, highlight the influence of task focus on the degree of coupling between the two modalities, and may have important implications for pronunciation training methods.

## Introduction

The nature of the link between speech perception (and comprehension) and speech production, and more specifically, the extent to which the two modalities rely on joint mechanisms and representations, is a central topic in language and speech science research. An area in which the characterization of this link is of particular interest is phonological development, in both early native (L1) learning and any second-language (L2) learning occurring later in life. This is because phonological development in a language requires that the listeners and speakers attune their perceptual systems to the phonological inventory of the language and also learn to execute the articulatory commands necessary to reproduce the phones in such inventory. In the case of late L2 learning, research has investigated the link between perception and production with respect to the acquisition of nonnative phonological categories in different ways, yet the largest body of research on this issue is probably the one

© The Author(s), 2023. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

examining the perception and production of particular L2 phones and subsequently assessing whether a relationship between individual performances in the two modalities can be found. Findings are mostly mixed, with several studies showing correlations, often of a moderate size and particularly for proficient L2 users (Flege, 1993; Flege et al., 1997; Melnik-Leroy et al., 2022), but some failing to do so (e.g., Peperkamp & Bouchon, 2011).

Recent endeavors to advance our understanding of this topic (Isbell, 2016; Nagle & Baese-Berk, 2022) have questioned the nature of the mixed findings outlined above, arguing that previous studies differ considerably in the tasks employed and, relatedly, in the types of processing they elicit. Crucially, Melnik-Leroy et al. (2022) provided compelling evidence that the alignment of the processing demands between the selected perception and production tasks modulates the strength of the relationship to be observed. They tested the perception and production of the French vowel contrast /u/-/y/ by English learners of French. Participants took part in a perception task assessing the discrimination of the two vowels in pseudowords and two production tasks, one in which they read pseudowords aloud and one that involved naming pictures of objects, which, unlike the other two tasks, entailed accessing stored lexical representations. Results revealed a robust link between perception and production when none of the two involved lexical access but not when they were mismatched in that respect.

All in all, the findings in Melnik-Leroy et al. (2022) very nicely subsume and validate the aforementioned concerns. However, the effects of processing level were only tested in one direction (i.e., prelexical perception vs. prelexical and lexical production). A critical remaining question is whether the same within-level association and the same dissociation in mismatching conditions emerge when the perception task itself relies heavily on lexical access. This is highly relevant because, within modalities, a recurrent finding has been that prelexical and lexical perception are not as tightly coupled as previously thought. First, it has been shown that accurate prelexical perception, while necessary, is not sufficient to guarantee accuracy in auditory spoken word recognition when the latter also depends on an accurate identification of particular L2 categories (e.g., Díaz et al., 2012; Llompарт, 2021b). Parallel findings have also been reported for production (Llompарт & Reinisch, 2019a). Secondly, in a similar way to what is observed across modalities, correlational analyses between prelexical and lexical perception have also rendered mixed results (e.g., Darcy & Holliday, 2019; Simonchyk & Darcy, 2017). Thus, although this needs not be taken to mean that there is no relationship between the two levels, such inconsistencies call for caution when considering whether results from prelexical tasks should generalize to lexical perception and production.

The present study set out to answer the question above by means of a joint analysis of part of the data in Llompарт and Reinisch (henceforth L&R; 2019a, 2019b). In L&R (2019b), a group of German learners of English of intermediate-to-high proficiency took part in an auditory lexical decision task where they were presented with real English words and nonwords containing specific consonant and vowel substitutions. The main contrast of interest was /ɛ/-/æ/, which is known to result in pervasive difficulties for this population because /æ/ is not part of the German vowel inventory and the two English vowels tend to be assimilated to German /ɛ/. L&R (2019a) tested the same set of participants on a phonetic categorization task, an imitation task in which the steps of the categorization continuum had to be imitated, and a word-reading task. The focus was again on /ɛ/-/æ/. The present study tests whether accurate perceptual identification of /ɛ/- and /æ/- (non)words in the lexical decision task can be predicted by

the production of /ɛ/ and /æ/ in word reading and by the production of the same vowels in the imitation task.

The auditory lexical decision task unambiguously requires lexical processing, as it prompts a comparison between a determined acoustic input and stored lexical representations with every decision. In contrast, the two production tasks are expected to diverge in this respect by virtue of their different (task) foci. Both word reading and imitation as implemented in L&R (2019a, 2019b) are controlled production tasks bound to trigger some level of focus on form in the learner (see Saito & Plonsky, 2019), and they both involve real word stimuli. However, the main focus in each task, as determined by the instructions, procedure, response types, and constraints, is critically different. As it will be further elaborated below, the word-reading task is focused on the holistic processing of English words, thus leading to lexical access and to some focus on meaning as well as form, whereas the imitation task focuses on the production of fine phonetic detail in the critical segments, which considerably limits the role to be played by lexical processing. Because of this, and building on the results of Melnik-Leroy et al. (2022), I hypothesize that only the word-reading production measure will relate to auditory word recognition in a meaningful way. Finally, it is worth acknowledging here that this hypothesis was not part of the original hypotheses of the previous studies. It was formulated and tested once the data had already been collected.

## Method

### *Participants*

Data from 34 of the participants (18 females, mean age = 25.21,  $SD = 4.35$ ) who took part in L&R (2019a, 2019b) were analyzed for the present study. These were the participants for which there was no missing data for any of the tasks. All participants were students at the Ludwig Maximilian University of Munich and participated in exchange of a small monetary compensation. The original recruitment criteria were that participants had not learned any language other than German before starting to learn English at school, had not spent more than 6 months in an English-speaking country, and were not enrolled in a language program at the university. All participants had given their informed consent to participate and all research outlined in the following sections was conducted in accordance with the Guidelines for Safeguarding Good Scientific Practice of the German Research Foundation and the Conference of Helsinki.

### *Materials and procedure*

Detailed descriptions of the methods and materials for the tasks reported here can be found in L&R (2019a, 2019b). In what follows, simplified descriptions will be provided to convey the main characteristics of the design. The tasks were administered in two sessions and in the order in which they are presented below. The lexical decision task was conducted in the first session and the word-reading and the imitation tasks between 1 and 3 weeks later in the second session.

#### *Perception: Lexical decision task*

The materials for the lexical decision task included 304 English words, 52 of which contained the vowels in the challenging L2 contrast /ɛ/-/æ/. Half of the words were selected to be presented as canonically produced (thus, as the real words in the task),

and the other half contained one phonological substitution that made them nonwords. For the critical contrast, this meant that 13 words with /æ/ appeared with the vowel canonically realized as [æ] and 13 different words were presented with /æ/ produced as [ɛ] (h[æ]mmer vs. \*dr[ɛ]gon). The same applied to items with /ɛ/ (d[ɛ]sert vs. \*l[æ]mon). A list of the /ɛ/-/æ/ stimuli is provided in [Appendix A](#).

All 304 words were recorded by a male speaker of Standard Southern British English. Participants were tested in a sound-attenuated booth at the university. The task was implemented in Psychopy 2 (Peirce et al., 2019). On each trial, two boxes were shown on the screen, a green one with “word” written on it on the left-hand side and a red one with “not a word” written on it on the right-hand side, and an auditory stimulus was presented over headphones at a comfortable listening level. Participants had to press “1” on the computer keyboard to indicate that the stimulus was a real word and “0” if they considered that the stimulus was not a real word. There was no time limit for participants’ responses. The 304 items were presented in a different randomized order for each participant.

#### *Production: Word-reading task*

The critical materials for the word-reading task were 13 English words with /ɛ/ and 13 words with /æ/. These were the same as the items included as real words in the lexical decision task. Thirteen words with /i/ and thirteen with /u/ were also included in the original study, but they can be considered fillers for the purposes of the current study. Note that these fillers also served to make the interest in /ɛ/ and /æ/ in this task less apparent. Participants were seated in front of a standing condenser microphone placed at approximately 30 cm from them in the sound-attenuated booth. They were informed that they would see English words appear on a computer screen and that their task was to read them aloud when instructed to do so.<sup>1</sup> They were always given some preview time to ensure that lexical access took place (Balota & Chumbley, 1985), and it was never mentioned that they should pay special attention to /ɛ/ and /æ/ or to vowel production in a more general sense. The words were presented only once and in a random order. Productions were sampled at 44.1 kHz with 16-bit quantization.

#### *Production: Imitation task*

For the imitation task, an 11-step *bet-bat* continuum was used. This continuum was created through duration manipulation and formant shifting by means of a Praat script (Boersma & Weenink, 2010). The durations and first and second formant values (F1 and F2) of the endpoints were taken from naturally produced tokens by the same talker who recorded the lexical decision stimuli. The steps in between were set to change linearly in all three dimensions (more details as well as endpoint values in L&R, 2019a). Participants were placed in exactly the same setting as for the word-reading task but were wearing over-ear headphones. They were told that, on every trial, they would hear a sequence of two stimuli and their task was to imitate the second of those as closely as possible right after the talker had finished producing it. The two stimuli on each trial were separated by 550 ms. The first stimulus was always one of the endpoints of the

<sup>1</sup> A word-familiarity questionnaire was used to ensure that participants knew the words in the study and were thus in the position to retrieve them from memory in the relevant tasks. They responded that they knew the meaning of 98.99% of the real words used for lexical decision and word reading and 100% of the words from which the nonwords in the lexical decision task were derived.

continuum, whereas the second could be any of the 11 steps (including the endpoints). The endpoint stimuli were presented as first stimulus to control for contrast effects from the immediately preceding trial. Each participant responded to a total of 88 *bet-bat* imitation trials, with each of the 11 steps being presented four times after each of the endpoints in a blocked manner. Participants' productions were recorded from 400 ms before the end of the audio file to 4 s after it. The next trial started automatically 1 s after the end of the recording.

As mentioned above, the imitation task is an interesting counterpart to the word-reading task in the context of this paper because, even though the endpoints of the continuum are two real English words, the main focus of the task is phonetic rather than lexical and its scope is mostly segmental. Only fine phonetic differences in the target vowels separate the different stimuli to be imitated, and, given that the vowel is the only segment that varies from trial to trial, participants are expected to allocate most of their attention to imitating the vowel after just a few trials. Furthermore, the very limited lexical variation in the stimuli, together with the fact that listeners do not necessarily know which word they are repeating on each trial and are not provided with any labels or feedback throughout, strongly suggests that the involvement of the lexicon during such a task is minimal. Note that similar arguments have been provided to question the involvement of the lexicon in perceptual categorization paradigms such as 2AFC tasks using real word stimuli (Amengual, 2016; Llompart, 2021b; Melnik & Peperkamp, 2021; see also Lively et al., 1993, for evidence of this).

### Data analysis and results

Following Melnik-Leroy et al. (2022), the relationship between perception and production was assessed by means of a series of generalized linear mixed-effects regression models on the perception (i.e., lexical decision) data. These models were run in R (version 3.6.3) using the lme4 package (version 1.1–23; Bates et al., 2015). Importantly, only nonword trials (mean % correct = 35.34,  $SD = 20.62$ ) were included in the analyses because participants were at ceiling with real word acceptance (mean % correct = 96.94,  $SD = 2.98$ ). The models had response (1 = correct, 0 = incorrect) as the binary dependent variable, and the predictors of interest included participant's individual scores for the word-reading and the imitation tasks and vowel (/ɛ/-/æ/). Vowel explained a considerable part of the variance in the original study (L&R, 2019b) and was included in order to reduce the risk of overestimating the effects of the production measures. The data set analyzed in this article and the code to reproduce the analyses reported are available at <https://osf.io/86wyc/>.

As production metrics for the word-reading and imitation tasks, Euclidean distances were calculated over (a) the individual means for the F1 and F2 of /ɛ/ and /æ/ in the corresponding words in the word-reading task and (b) the individual means of the F1 and F2 of /ɛ/ and /æ/ in the imitations of only the endpoint steps of the continuum (i.e., Steps 1 and 11) in the imitation task. This is because the endpoint steps were the only ones that had the original F1, F2, and duration values by the native speaker (see L&R, 2019a, for more details). The mean Euclidean distance between the two vowels in the word-reading task for the present sample was 124 Hz ( $SD = 62$ ), and the mean distance for the imitated endpoints in the imitation task was 370 Hz ( $SD = 123$ ). The Euclidean distance for one participant in the word-reading task (340 Hz) was an outlier in the current distribution of values, in which very little separation between the two categories was the norm. However, the raw F1 and F2 values of that speaker did not

**Table 1.** Correlation matrix for individual measures in the three tasks compared in this study

	Lexical decision	Word reading	Imitation
Lexical decision	1.00		
Word reading	.43*	1.00	
Imitation	.19	.30	1.00

Note. \* $p < .05$ .

point toward any abnormality, and they actually approximated those typical of native realizations of the contrast (e.g., the values of the native talker). Because of this, data from this participant were kept for all tasks but their individual metric for the word-reading task was substituted for the second-highest value in the sample (241 Hz) to avoid an undue influence of the outlier on the results.<sup>2</sup> The individual numeric variables for word reading and imitation were centered and scaled using the `scale()` function in R before entering the analyses. A correlation matrix summarizing the relationships between individual measures for each of the three tasks is presented in Table 1. For the lexical decision task, the by-participant measure was the proportion of correct / $\varepsilon$ /- and / $\text{æ}$ /-nonword rejections.

Model comparisons and model selection followed the step-wise procedure described in Melnik-Leroy et al. (2022). A base model was created first that included only an intercept and random intercepts for participant and item and a random slope for vowel over participant. Subsequently, at each step, it was tested whether each predictor improved the model's fit on its own by means of likelihood-ratio tests. Then the effect with the lowest significant  $p$  value was added to the model and the procedure was repeated. Once more than one predictor was added to the model, the interaction between them was tested. The final best-fitting model included main effects of vowel and word reading and no effect of imitation or any significant interaction (see Table 2).

The effect of vowel indicates that accuracy was higher for nonwords in which / $\varepsilon$ / had been substituted by [æ], like \*l[æ]mon (mean % correct = 47.06,  $SD = 24.16$ ) than vice versa (e.g., \*dr[ $\varepsilon$ ]gon; mean % correct = 23.38,  $SD = 22.45$ ). The effect of word reading shows that larger distances between the two vowels in the word-reading task predicted higher accuracy in nonword rejection for items with these vowels in the lexical decision task. This aligns with the significant correlation reported in Table 1. In addition, the

**Table 2.** Coefficients and results of log-likelihood comparisons for each retained effect in the final best-fitting model. The results of log-likelihood comparisons between the best-fitting model and two separate models additionally containing imitation and the interaction between vowel and word reading, respectively, are also provided

Predictor	$\beta$	SE	$z$	$\chi^2$	df	$p$
Intercept	-1.75	0.44	-4.03	–	–	–
Vowel	1.58	0.55	2.88	7.40	1	< .01
Word reading	0.61	0.22	2.80	7.24	1	< .01
Imitation	–	–	–	0.23	1	.63
Vowel $\times$ Word reading	–	–	–	0.07	1	.78

<sup>2</sup>It was assessed whether the analyses rendered different results if the outlier was not replaced and they did not.

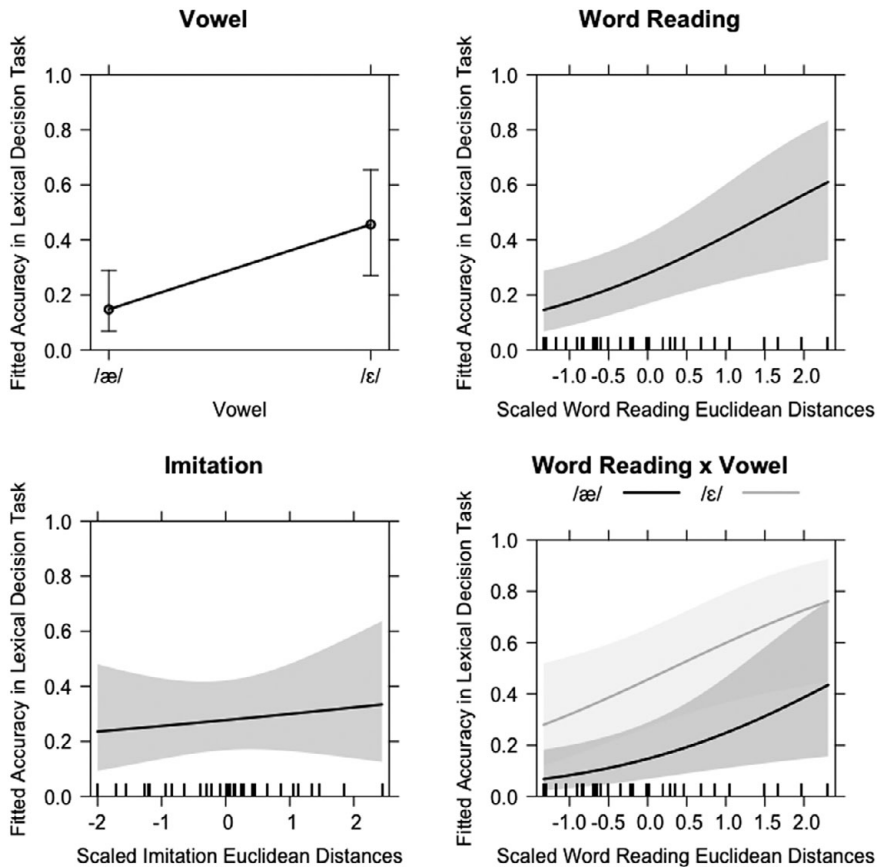


Figure 1. Effect plots depicting the fitted probability of accurate nonword rejection in the lexical decision task as a function of vowel (top left panel), word reading (top right panel), imitation (bottom left panel), and the interaction between word reading and vowel (bottom right panel).

absence of an interaction between vowel and word reading indicates that this relationship was not qualified by target vowel. Finally, the analyses provided no evidence that vowel distance in the imitation task predicted accuracy in the lexical decision task. Effect plots for vowel and word reading in the best-fitting model as well as for imitation and the interaction between word reading and vowel when these were separately added to the best-fitting model were obtained using the `predictorEffect()` function of the `effects` package (version 4.2–2; Fox & Weisberg, 2018) and are provided in Figure 1.

## Discussion

The present study sought to investigate whether a link between the perception and the production of a challenging L2 vowel contrast could be found when both were assessed through tasks relying on lexical processing. Most research in this area to date has focused exclusively on either prelexical processing (Jia et al., 2006) or on a mixture of lexical and prelexical processing (Peperkamp & Bouchon, 2011) and, until recently

(Isbell, 2016; Melnik-Leroy et al., 2022; Nagle & Baese-Berk, 2022), little attention had been given to the potential modulating effects that task focus and processing level may have on the perception–production link. Examining this link in tasks that involve lexical access is critical not only because of the mismatches observed between prelexical and lexical tasks within each of the two modalities discussed in the Introduction, but also because it has been shown that lexical access is influenced by the learners' lexical knowledge (e.g., vocabulary size and depth) and by the lexical characteristics of individual L2 words (Daidone & Darcy, 2021; Llompart, 2021a). These additional interacting factors could have an influence on the coupling between the perceptual and productive mastery of challenging nonnative contrasts that learners display at the lexical level versus in tasks devoid or quasi-devoid of lexical processing.

Nonetheless, the main finding of this study was that, when following a very similar analysis procedure to that in Melnik-Leroy et al. (2022), individual Euclidean distances between /ɛ/ and /æ/ in a word-reading task indeed predicted the extent to which L2 learners were able to reject nonwords created by swapping the two vowels (e.g., \*l[æ]mon, \*dr[ɛ]gon) in an auditory lexical decision task. This provides further evidence in favor of a link between perception and production in L2 speech learning, which is well in line with the postulates of the major L2-speech learning models, such as the perceptual assimilation model-L2 (Best & Tyler, 2007), which predicts a tight connection between modalities, and both the original speech learning model (Flege, 1995) and its recently revised version (Flege & Bohn, 2021), which envision a flexible yet apparent relationship between perception and production. Furthermore, the present results crucially extend previous findings to perception and production at a lexical level as examined through tasks without an inherent focus on accurately perceiving and producing acoustic differences in particular L2 phones.

Such an extension to lexical processing tasks is remarkable for several reasons: First, very much related to the argument above, neither the lexical decision task nor the word-reading task used here asked participants to provide any direct responses about the critical vowels. Instead, they both encouraged that the stimuli were evaluated (in perception) and constructed (in production) holistically as word units. Besides, these units could contain the target vowels or not, as there were fillers without them in the two tasks. Second, the relationship illustrated in Figure 1 (top right panel) arose even though individual performances in the two tasks were generally rather poor. In spite of them being of an intermediate-to-high proficiency, the L2 learners as a group were actually below chance in the perception task (35% correct) and the average Euclidean distance between the vowels in the production task was rather small (124 Hz), especially compared with that in imitation (370 Hz). Finally, production predicted perception despite the fact that ensuring that the tasks required lexical access involved anticipating the influence of a variety of lexicon-related factors such as L2 vocabulary size and the lexical characteristics of the words (e.g., frequency), which are expected to have a weaker influence on prelexical processing (but see Bundgaard-Nielsen et al., 2011). All in all, the present results suggest that the perception–production link, when assessed while taking task focus and level of processing into account, is quite robust, as there is evidence for it even when learners (a) are not guided to allocate special attention to the target contrast, (b) are as a group shown to struggle to differentiate between the target phones in both modalities, and (c) are susceptible to higher level influences on their accomplishment.

Conversely to the word-reading task, individual Euclidean distances between the vowels of the imitated endpoint steps did not predict lexical decision accuracy. Therefore, these results mirror those of Melnik-Leroy et al. (2022) in that a relationship



between perception and production could be found when the two were measured in tasks with similar affordances but not when they were less symmetrical in that respect. Of course, the lack of an effect for imitation should be interpreted with caution, as it is possible that the study did not have enough power to detect it (Lakens, 2022). However, Figure 1 (bottom left panel) suggests that, if any, the contribution of imitation scores to explaining perception performance in the present data set is certainly limited.

The inability to find a relationship between lexical decision and imitation is particularly interesting for at least three reasons. First, any imitation task has an obvious speech perception component: participants first need to perceive the stimuli to be able to imitate them. Therefore, such need for perception would in principle bring imitated productions closer to any perception measure than productions elicited otherwise (see Kato and Baese-Berk, 2020). Second, both word reading and imitation involved real word stimuli. Thus, if the crucial requirement for a perception–production link to be found was that the stimuli shared the same lexical status, a relationship with lexical decision task accuracy should have been observable for imitation just as it was for word reading. In third place, in L&R (2019a) another measure derived from the same imitation task was found to correlate with an individual measure of perceptual categorization in a 2AFC task sharing the same phonetic and segmental focus. Particularly considering these last two points, the results very strongly suggest that it is the focus of each task and how this in turn determines the extent to which lexical processing is required that is a primary modulating factor of the relationship between L2 perception and production.

It is true, however, that because of the design of the original studies and the tasks therein, other alternatives should also be entertained, like the possibility that the findings partly stem from (a) order effects and asymmetric fatigue effects and/or (b) differences across tasks in terms of the lexical and phonetic contexts in which the critical vowels appeared. Regarding (a), that order or fatigue played a major role seems rather unlikely. Concerning order effects, the perception task was almost equally removed in time for both production tasks, and it is hard to envisage how, even if the production tasks were always completed in the same order, the word-reading task may have influenced the imitation task, as the interest in /ɛ/ and /æ/ was much less obvious in the former than the latter. Potential fatigue effects are similarly improbable, as the word-reading task took only around 5 min.

With respect to (b) above, lexical decision and word reading did indeed present more variation in words and phonetic contexts for the critical vowels than the imitation task. Although it is unclear that increased variability can be considered an advantage for a relationship between the former two tasks to arise, given that said contexts were completely different (see Appendix A), the constraints of the imitation task might have hindered to some extent that a relationship with perception is observed for this task. This is because the repetitive imitation setting could have provided an opportunity for participants to modify (and perfect) their articulation of the target vowels along the task in a way that was not possible in the other tasks. Thus, if one assumes that not all participants benefitted from this repetition to the same extent, this may have contributed to making the lexical perception task and the segmentally focused imitation task more dissimilar. It is thus advisable for future research to take these aspects into account when devising the cross-modal comparisons of interest.

All in all, even though the above-mentioned confounding factors cannot be fully ruled out, I believe that task focus effects still appear to be the most likely cause of the patterns observed: when the perception and production task both encourage holistic processing and lexical access, the link between the two is tighter than when production

has a stronger focus on phonetic detail regarding the critical segments. This claim aligns well with the distinction between the *phonological* and the *phonetic* mode made in the automatic selective perception model (Strange, 2011; Strange & Schaffer, 2008). The phonological mode is automated and does not require focused attention, and it is therefore thought to be active when accessing word meanings is required, as in a lexical decision task. The phonetic mode, on the contrary, is concerned with accessing context-dependent phonetic detail and consequently requires selective attention and sufficient cognitive resources, like for example in categorization tasks. Considering this, a parallel can be established for the two speech-production tasks used here, with the word-reading task prompting the engagement of a phonological (or phonolexical) processing mode—just as the auditory lexical decision task assessing perception—and the imitation task relying mostly on phonetic processing.

Finally, although the relationship between the perception and production of challenging L2 phonological distinctions in lexical tasks needs further study (e.g., by targeting different segment types and also suprasegmental contrastive features), the present results could potentially have important pedagogical implications. Communication involves building meaning out of words and (generally) not isolated sounds, and training and teaching methods seeking to aid in the development of a nonnative phonological inventory should be aware of the ubiquitous mismatches to be found between prelexical and lexical processing and of how these relate to task demands. Therefore, if, at least at higher levels of proficiency, a major goal of pronunciation training is that L2 learners improve their production of particular L2 phones when they focus on conveying the meaning of words in real communication, perceptual training paradigms with a strong lexical component (e.g., “feedbacked” lexical decision tasks and picture-word matching tasks) may be more effective than the phonetic training procedures relying on categorization and reduced sets of (non)words that are now the norm. This is an empirical question at the moment, yet it is one that can hopefully set the stage for much exciting work on the perception–production link in L2 speech learning in the future.

**Acknowledgments.** Part of this work was conducted while the author was supported by the Alexander von Humboldt Professorship (ID-1195918) awarded to Ewa Dąbrowska, Chair of Language and Cognition at the Friedrich Alexander University Erlangen-Nuremberg. I would like to thank the four anonymous reviewers and the editor of this article for their very helpful suggestions.

**Data availability statement.** The experiment in this article earned the Open Data badge for transparent practices. The data are available at: <https://osf.io/86wyc/>.

**Competing interest.** The author declares none.

## References

- Amengual, M. (2016). The perception of language-specific phonetic categories does not guarantee accurate phonological representations in the lexicon of early bilinguals. *Applied Psycholinguistics*, 37, 1221–1251.
- Balota, D. A., & Chumbley, J. I. (1985). The locus of word-frequency effects in the pronunciation task: Lexical access and/or production? *Journal of Memory and Language*, 24, 89–106.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48.
- Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O. S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 13–34). John Benjamins.

- Boersma, P., & Weenink, D. (2010). Praat: doing phonetics by computer. Version 5.4.22, <http://www.praat.org/>
- Bundgaard-Nielsen, R. L., Best, C. T., & Tyler, M. D. (2011). Vocabulary size is associated with second-language vowel perception performance in adult learners. *Studies in Second Language Acquisition*, 33, 433–461.
- Daidone, D., & Darcy, I. (2021). Vocabulary size is a key factor in predicting second-language lexical encoding accuracy. *Frontiers in Psychology*, 12, Article 2769.
- Darcy, I., & Holliday, J. J. (2019). Teaching an old word new tricks: Phonological updates in the L2 mental lexicon. In J. Levis, C. Nagle, & E. Today (Eds.), *Proceedings of the 10th Pronunciation in Second Language Learning and Teaching Conference* (pp. 10–26). Iowa State University.
- Díaz, B., Mitterer, H., Broersma, M., & Sebastián-Gallés, N. (2012). Individual differences in late bilinguals' L2 phonological processes: From acoustic-phonetic analysis to lexical access. *Learning and Individual Differences*, 22, 680–689.
- Flege, J. E. (1993). Production and perception of a novel, second-language phonetic contrast. *The Journal of the Acoustical Society of America*, 93, 1589–1608.
- Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, 92, 233–277.
- Flege, J. E., & Bohn, O. S. (2021). The revised speech learning model (SLM-r). In R. Wayland (Ed.), *Second language speech learning: Theoretical and empirical progress* (3–83). Cambridge University Press.
- Flege, J. E., Bohn, O.-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, 25, 437–470.
- Fox, J., & Weisberg, S. (2018). Visualizing fit and lack of fit in complex regression models with predictor effect plots and partial residuals. *Journal of Statistical Software*, 87, 1–27.
- Isbell, D. R. (2016). The perception-production link in L2 phonology. *MSU Working Papers in Second Language Studies*, 7, 57–67.
- Jia, G., Strange, W., Wu, Y., Collado, J., & Guan, Q. (2006). Perception and production of English vowels by Mandarin speakers: Age-related differences vary with amount of L2 exposure. *The Journal of the Acoustical Society of America*, 119, 1118–1130.
- Kato, M., & Baese-Berk, M. M. (2020). The effect of input prompts on the relationship between perception and production of non-native sounds. *Journal of Phonetics*, 79, Article 100964.
- Lakens, D. (2022). Sample size justification. *Collabra: Psychology*, 8, Article 33267.
- Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *The Journal of the Acoustical Society of America*, 94, 1242–1255.
- Llompart, M. (2021a). Lexical and phonetic influences on the phonolexical encoding of difficult second-language contrasts: insights from nonword rejection. *Frontiers in Psychology*, 12, Article 659852.
- Llompart, M. (2021b). Phonetic categorization ability and vocabulary size contribute to the encoding of difficult second-language phonological contrasts into the lexicon. *Bilingualism: Language and Cognition*, 24, 481–496.
- Llompart, M., & Reinisch, E. (2019a). Imitation in a second language relies on phonological categories but does not reflect the productive usage of difficult sound contrasts. *Language and Speech*, 62, 594–622.
- Llompart, M., & Reinisch, E. (2019b). Robustness of phonolexical representations relates to phonetic flexibility for difficult second language sound contrasts. *Bilingualism: Language and Cognition*, 22, 1085–1100.
- Melnik, G. A., & Peperkamp, S. (2021). High-variability phonetic training enhances second language lexical processing: Evidence from online training of French learners of English. *Bilingualism: Language and Cognition*, 24, 497–506.
- Melnik-Leroy, G. A., Turnbull, R., & Peperkamp, S. (2022). On the relationship between perception and production of L2 sounds: Evidence from Anglophones' processing of the French /u/-/y/ contrast. *Second Language Research*, 38, 581–605.
- Nagle, C. L., & Baese-Berk, M. M. (2022). Advancing the state of the art in L2 speech perception-production research: Revisiting theoretical assumptions and methodological practices. *Studies in Second Language Acquisition*, 44, 580–605.
- Pearce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 51, 195–203.

- Peperkamp, S., & Bouchon, C. (2011, August 27–31). *The relation between perception and production in L2 phonological processing* [Paper presentation]. Twelfth Annual Conference of the International Speech Communication Association, Florence, Italy.
- Saito, K., & Plonsky, L. (2019). Effects of second language pronunciation teaching revisited: A proposed measurement framework and meta-analysis. *Language Learning*, 69, 652–708.
- Simonchyk, A., & Darcy, I. (2017). Lexical encoding and perception of palatalized consonants in L2 Russian. In M. O'Brien & J. Levis (Eds.), *Proceedings of the 8th Pronunciation in Second Language Learning and Teaching Conference* (pp. 121–132). Iowa State University.
- Strange, W. (2011). Automatic selective perception (ASP) of first and second language speech: A working model. *Journal of Phonetics*, 39, 456–466.
- Strange, W., & Shafer, V. L. (2008). Speech perception in second language learners: The re-education of selective perception. In M. Zampini & J. Hansen (Eds.), *Phonology and second language acquisition* (pp. 153–191). Cambridge University Press.

## Appendix A

Words and nonwords used in the lexical decision task (Llompart & Reinisch, 2019b). For the nonwords, the real word from which they were derived is also provided. The words were also used as stimuli in the word-reading task (Llompart and Reinisch, 2019a).

Words	Nonwords
centre	ch[æ]rry; cherry
check	ch[æ]ss; chess
chest	d[æ]sk; desk
death	dr[æ]ss; dress
dentist	fr[æ]sh; fresh
desert	h[æ]th; health
helmet	h[æ]lp; help
helpful	l[æ]gend; legend
lecture	l[æ]mon; lemon
record	l[æ]sson; lesson
rest	s[æ]ntence; sentence
smell	w[æ]ther; weather
special	y[æ]llow; yellow
champion	ch[ɛ]nnel; channel
damage	dr[ɛ]gon; dragon
fact	f[ɛ]ctor; factor
fashion	f[ɛ]ctory; factory
hammer	fl[ɛ]g; flag
hand	g[ɛ]llery; gallery
happy	h[ɛ]bit; habit
language	l[ɛ]mp; lamp
national	r[ɛ]mp; ramp
Saturday	sc[ɛ]ndal; scandal
shadow	sp[ɛ]sh; splash
smash	st[ɛ]ndard; standard
snack	th[ɛ]nk; thank

**Cite this article:** Llompart, M. (2024). On the effects of task focus and processing level on the perception–production link in second-language speech learning. *Studies in Second Language Acquisition*, 46: 214–226. <https://doi.org/10.1017/S0272263123000414>