



The Impact of Free Allophonic Variation on the Perception of Second Language Phonological Categories

Eva Reinisch^{1,2*}, Katharina I. Juhl¹ and Miquel Llompart³

¹ Institute of Phonetics and Speech Processing, Ludwig Maximilian University, Munich, Germany, ² Acoustics Research Institute, Austrian Academy of Sciences, Vienna, Austria, ³ Department of English and American Studies, Friedrich Alexander University, Erlangen, Germany

OPEN ACCESS

Edited by:

Martijn Baart,
Tilburg University, Netherlands

Reviewed by:

Willy Serniclaes,
Université Paris Descartes, France
Karen Mulak,
University of Maryland, College Park,
United States

*Correspondence:

Eva Reinisch
ereinisch@kfs.oew.ac.at

Specialty section:

This article was submitted to
Language Sciences,
a section of the journal
Frontiers in Communication

Received: 20 January 2020

Accepted: 10 June 2020

Published: 29 July 2020

Citation:

Reinisch E, Juhl KI and Llompart M
(2020) The Impact of Free Allophonic
Variation on the Perception of Second
Language Phonological Categories.
Front. Commun. 5:47.
doi: 10.3389/fcomm.2020.00047

When learning the phonological categories of a second language (L2), learners have to deal with phonetic variation. For instance, allophonic variant forms have to be recognized as the same phoneme. A minimal pair identification task was used to assess how late Spanish learners of German perceive the phonological /r/-/h/ contrast. German /r/ was realized as one of three allophones ([r], [ʀ], [ʁ]) that vary in phonetic similarity to /h/ as well as to the typical phonetic form of Spanish /r/ (i.e., [r]). Results showed that learners were very good at identifying all German variant forms (>90% correct). However, [ʁ], which is phonetically closest to German /h/ and furthest from Spanish /r/, was identified significantly worse than [r] and [ʀ]. Relating these results to a cross-language perception task where learners were asked to map the German allophones of /r/ and the phoneme /h/ to different L1 phonological categories further showed that those learners were best at identifying words with [ʁ] who consistently matched it to a single L1 category. Surprisingly this category did not always have to be the phonologically matching Spanish /r/. We conclude that phonological and phonetic relations between the learners' L1 and L2 are important in identifying allophones of the same L2 category.

Keywords: second language learning, speech perception, sound category learning, free allophones, German, Spanish

INTRODUCTION

One of the fundamental issues in second language (L2) learning is to acquire a new phonological system. This entails for the learner to grasp what types of acoustic and articulatory differences between phones lead to differences in word meaning in the L2 independently of whether these differences would have the same effects in their native language (L1). However, the L1 phonological system has been shown to have a major influence on L2 learning leading to major difficulties in the mastery for at least a subset of L2 phonological contrasts. These difficulties are commonly attributed to the fact that, at the initial stages of L2 learning, the L1 acts like a sieve (Trubetzkoy, 1977; Flege, 1995; Best and Tyler, 2007), modulating how L2 phones are perceived and produced. Importantly, L2 learners need to establish a non-native phonological system whilst facing the vast variability that characterizes speech. This variability stems from many different sources, among others, from physiological differences between speakers, regional, social or non-native accents of speakers or even idiosyncratic variation. In the present study, we investigate how Spanish learners

of German *deal*, in perception, with variability caused by different allophonic variants of an L2 phonological category that are typically attributed to a combination of regional and speakers' idiosyncratic variation.

Variability within a given language's phonological system is quite common. That is, certain phonemes of the language are realized as different variant forms (i.e., allophones). The occurrence of these variants may depend, for instance, on their position in the word or syllable. Examples for positional allophones are the Dutch /r/ realized in an "English-like" manner as [ɾ] in syllable-final position (vs. the default variants [r] or [ʁ]; Van Bezooijen, 2005), or phonologically voiced obstruents that are restricted to non-final position in many languages (e.g., Dutch, German, Polish; i.e., in final position they are devoiced). Native listeners have been shown to use position-specific processing strategies for these types of allophones (Mitterer et al., 2013, 2018; Mitterer and Reinisch, 2017). Importantly, regional and idiosyncratic variation between speakers may lead to the presence of allophones that appear in free variation, that is, without positional restrictions. A case in point are the allophones of the German rhotic, which are the focus of the present study.

German /r/ has many different variants that radically differ from each other in terms of their acoustics and articulation (e.g., Ulbrich and Ulbrich, 2007). Here we focus on three main variants occurring in syllable onsets¹ that are listed as allophones in phonological descriptions (e.g., Wiese, 1996): the alveolar trill [r], the uvular trill [R] and the uvular fricative [ʁ]. The trills are phonetically characterized by 1–3 brief amplitude minima caused by the occlusion of the airflow by the tongue tip (for [r]) or the tongue dorsum (for [R]). The amplitude minima tend to be more regular for the alveolar trill due to the greater flexibility of the tongue tip. The uvular fricative [ʁ] is frequently considered to be produced as an approximant with less constriction of the airflow and hence a lower amplitude spectrum and less stochastic noise in the signal than is typical for a fricative (Kohler, 1995; Wiese, 1996). However, we use "uvular fricative" to refer to one allophone that includes approximant and fricative realizations but is crucially different in acoustics and articulation from the two trilled allophones. All allophones may be subject to further variation such that they may be (partially) devoiced in voiceless phonetic contexts (cf. Ulbrich and Ulbrich, 2007). Importantly, which of the three allophones under consideration here is used in speech production does not depend on the immediate phonetic context or position in a word but varies between (and to some extent even within) speakers as well as between regions (Wiese, 1996, 2003). All allophones are encountered in Southern Germany, where the present study took place, even though [ʁ], due to its association with Standard German and prevalence in national media, is likely the most frequently heard variant.

In an immersive L2 learning scenario, allophonic variation for German /r/ means that learners are exposed to all variants.

They therefore need to learn that speakers may differ in their realization of /r/ but that this does not affect the meaning of words. That is, they have to learn to recognize each of these variant forms as referring to the same phoneme category. Given the assumption of L2 sound learning models that, at least at initial stages of learning, L2 phones are perceived in relation to the L1 phonological inventory (e.g., Flege, 1995; Best and Tyler, 2007; van Leussen and Escudero, 2015), the question arises as to how learners of German would deal with the allophonic variation described above if their L1 shared a phoneme category /r/ but lacked this specific allophonic variation, as is the case for native speakers of Spanish who are learning German. Therefore, in the present study, we addressed how Spanish learners of German perceive the German phonological category /r/ and its allophonic variant forms. For these learners, the allophones of German /r/ all match the /r/ of their L1 at the phonological level but crucially differ in their degree of phonetic match with the typical phonetic characteristics of that L1 phoneme. Note that, in order to highlight the cross-language match of the phonological category, throughout this paper the notation /r/ will be used to refer to the phoneme in both languages. This is despite the fact that the standard and most common variant in current Standard German² pronunciation could be considered the uvular fricative (hence /ʁ/; though the notation /r/ is common in phonological descriptions, e.g., Wiese, 2003).

Given that Spanish is spoken in many different countries and is therefore subject to regional variation (including some variation in the production of /r/; see Hualde, 2005, for a comprehensive summary), we restricted our learner set to speakers of Standard Iberian (Castilian) Spanish. Castilian Spanish has two types of rhotics that are phonemically contrastive in word-medial position: a trill (/r/) and a tap (/ɾ/) which are both produced at an alveolar place of articulation with multiple apical closures for the trill and one brief apical closure for the tap (Hualde et al., 2010). In terms of acoustics, both sounds are voiced, with a lower overall amplitude than the surrounding vowels. The acoustic difference between trill and tap is then reflected in overall duration (with the trill being longer) as well as the number of brief regular drops in amplitude due to airflow constriction (i.e., one for the tap, and up to 5 for the trill; Navarro-Tomás, 1916; Blecua, 2001; Hualde et al., 2010).

Importantly, the allophones of German /r/ match the acoustic and articulatory characteristics of a typical Spanish /r/ to different extents. The German alveolar trill [r] is a close match to the typical realization of Spanish /r/, even though the number of occlusions during the trill may be fewer and duration hence shorter in German. The German uvular trill [R] matches Spanish /r/ in manner of articulation (i.e., trill) but differs from it in place of articulation. In terms of acoustics, the presence of amplitude minima due to the brief occlusions of the airflow makes [R] similar to Spanish /r/ though maybe not as good a match as the alveolar trill. Finally, the third German allophonic variant, the uvular fricative [ʁ], does not match Spanish /r/ in either

¹The present study focuses on variants in syllable onset position. In syllable coda, in postvocalic position, additional vocalized variants of the rhotic are common that will not be discussed here. The interested reader is referred to Kohler (1995, chapter 6.1.16), Ulbrich and Ulbrich (2007) and Wiese (1996, chapters 6.3.1 and 7.4.4) for details.

²For detailed discussions of the historical development and regional distributions of the German rhotics see, for instance, Kohler (1995), Schiller (1999), Ulbrich and Ulbrich (2007), or Wiese (1996, 2003).

manner or place of articulation. In terms of acoustics, it lacks the amplitude drops associated with the airflow occlusions in the trills but instead shows a higher proportion of noise over the spectrum.

The main question addressed in this study is whether Spanish learners of German perceive the three allophones of German /r/ differently depending on their acoustic and articulatory (i.e., phonetic) match to the typical realization of Spanish /r/. Learners are expected to know about the existence of the German phoneme /r/ and its functional equivalence to Spanish /r/. In both languages the phoneme is spelled with the letter <r> and, since both are Indo-European languages, they share (near) cognates such as German *Rose*, Spanish *rosa* (English “rose”). Nevertheless, the differences in acoustics and articulation between each of the German allophones and the typical phonetic characteristics of Spanish /r/ may result in differences in ease of recognition—even if learners are familiar with all three German allophones. Testing the perception of this set of L2 allophones will therefore contribute to elucidating to what degree L2 phoneme perception is influenced by the phonological and phonetic mappings of phones between and within the two languages of the learners.

In order to make predictions about the potential differences in difficulty of recognizing the different German allophones, we refer to models of second language learning such as the Speech Learning Model (SLM; Flege, 1995), the Perceptual Assimilation Model for L2 learning (PAM-L2; Best and Tyler, 2007) or the Second Language Linguistic Perception (L2LP) model (van Leussen and Escudero, 2015). Generalizing across details, in these models, L2 phones that are phonetically³ similar to L1 phonemes are assumed to be (perceptually) assimilated to these L1 categories diminishing the chances of acquiring these as separate L2 categories. That is, L1 and L2 phones that are phonetically sufficiently close will be established as one merged L1-L2 phoneme category. If an L2 phone is not (straightforwardly) assimilated to an L1 phoneme, a new L2 phoneme for this phone can be established. The question now is what perceptual patterns could be expected if the L1 and L2 share a phonological category but some of the L2 allophones of this phoneme are more similar to the typical realization of the L1 phoneme than others. Since the goal of our study was to test L2 learners’ behavior rather than to compare specific assumptions of the different models, we picked the terminology of one, PAM-L2, to base our predictions on. This is, of course, not to say that other models could not be used to make predictions, but the fact that PAM-L2 specifically focuses on the perception of L2 phoneme contrasts and studies building on this model often make use of cross-language perception tasks (which we also used here) contributed to this choice.

According to PAM-L2, learners should detect that German and Spanish /r/ are functionally equivalent categories at the

phonological level. The model further assumes that if the typical phonetic realization of the L2 phoneme sufficiently differs between languages (such as e.g., English /ɹ/ vs. French /ʁ/, the example given in Best and Tyler, 2007, where both phones refer to the functionally equivalent phoneme of a rhotic, spelled <r> in both languages) then learners should be able to accommodate a phonetic sub-category for this L2 phone within their L1 phoneme category. Note that this example from English vs. French matches our case of German [ʁ] being phonologically but not phonetically similar to Spanish /r/. Based on phonetic distance, learners would hence be predicted to be able to establish a separate phonetic (sub)-category [ʁ] that is part of the functionally matching L1-L2 phonological category /r/.

However, the present case for Spanish learners of German is somewhat more complex, as there are three different L2 allophones that have to be functionally mapped onto the same phonological category. Importantly, these L2 allophones differ not only in their degree of phonetic match to the typical realization of the functionally equivalent L1 phoneme (Spanish /r/), but also in how phonetically similar they are to other, functionally-distinct L1 and L2 phonemes. For this reason, we specifically assessed learners’ identification of the three allophones of German /r/ vs. another German phoneme, /h/. This is because one of the German allophones of /r/, namely the velar fricative [ʁ], is phonetically similar to German /h/ and, as such, it might lead to perceptual confusion between the German phonemes /r/ and /h/ for Spanish learners.

The glottal fricative /h/ in German is variably produced as voiced or voiceless, is restricted to syllable initial position, and tends to show weak frication with formant-like patterns matching the formants of adjacent vowels. This phoneme is lacking in the Spanish phonological inventory, with the phonetically closest phoneme being /x/, the velar fricative, which typically shows stronger frication than German /h/ (Hualde et al., 2010; note that German [x] as allophone of /ç/ does not occur in syllable-onset, Wiese, 1996). Critically, due to the lack of /h/ in Spanish and its likely perception as Spanish /x/, L2 learners of German could perceive /h/ as phonetically similar to the German uvular fricative variant of /r/, which shares acoustic and articulatory characteristics with German /h/ as well as Spanish /x/.

Building from this, we asked Spanish learners of German to identify German minimal word pairs starting with /r/ and /h/ (*Rose* “rose”–*Hose* “pants”) where /r/ could be produced as either of the variants [r], [R], or [ʁ]. Identification was preferred over a discrimination task and real words were chosen over non-words to highlight the functional difference between the phoneme categories. Yet, since, as argued above, based on phonetic characteristics [ʁ] may be perceptually confusable with /h/, an identification task allows for the assessment of the relation between phonological and phonetic influences in L2 sound/word identification. If learners easily recognize [ʁ] as an allophone of the German phoneme /r/ then little confusion with the phoneme /h/ is expected. However, if phonetic similarity to one or more phoneme categories in the L2 (or the L1) plays a role in L2 phoneme identification, then learners should correctly recognize the intended word less often when it had been produced with [ʁ] than the two trilled allophones of German /r/ that are not

³PAM (Best, 1994, 1995) and PAM-L2 (Best and Tyler, 2007) specifically refer to the articulatory similarity between sounds as the model is embedded in the direct realist framework of perception (Gibson, 1966, 1991). SLM (Flege, 1995) and the L2LP model (van Leussen and Escudero, 2015) base their predictions on acoustics. However, since in many cases, including the present study, articulatory vs. acoustic differences would not lead to different predictions, we will use the term “phonetic” to refer to articulatory and/or acoustic differences between sounds.

TABLE 1 | Summary of most likely cross-language perception patterns for Spanish learners of German based on phonological match (left) and phonetic match based on manner of articulation (right).

Phonological match		Phonetic match			
German	Spanish	German		Spanish	
/r/ (all allophones)	↔	/r/	[r]	↔	[r] [ɾ]
/h/	↔	/x/	[R]	↔	[r] [ɾ]
			[ʁ] * [h]	↔	[x] [ɣ]

Note that additional phonetic associations could be expected based on place of articulation for the trilled variants, however, as discussed in the main text, the trilled manner of articulation likely represents the more salient perceptual characteristic. Potential confusions between German phones (i.e., confusions between L2 phones) based on their respective acoustic correlates as described in the main text are indicated by *.

phonetically similar to any other German phoneme (but are crucially similar to Spanish /r/).

Since L2 learners are only in the course of establishing an L2 phonological system, it is likely that the L1 phonological inventory plays a role in L2 perception even for medium-to-advanced L2 learners—in terms of phonetic as well as phonological match to the L2 categories (e.g., Díaz et al., 2012). Therefore, another aim of the present study was to assess whether individual learners' performance in identifying the German allophones [r], [R], and [ʁ] as the phoneme /r/ relates to their assessment of how well these allophones match phonetically similar phonemes of their L1. To address this issue, learners were asked to perform a cross-language perception task prior to the L2 minimal pair identification task. The aim of the cross-language perception task was to test whether the strength of the association of German /r/ with the phonologically matching Spanish category /r/ would differ between the three different German allophones of /r/, and whether other Spanish phonemes (i.e., /r/, /ɾ/, /k/, /x/, and /g/'s allophone [ɣ]) would be considered possible L1 alternatives for (some of) the German allophones—and if so, which ones and how many of them.

In terms of phonetic similarity, we would expect that German [r] is straightforwardly mapped to Spanish /r/ with maybe a few /r/-responses due to the phonetic differences between the German and Spanish trills discussed above. The German uvular trill [R] is also predicted to be mainly mapped to Spanish /r/. However, since in terms of place of articulation (though not manner or acoustics) the closest consonants to [R] in Spanish are the velar stops /k/ and /g/ and the velar fricative /x/, German [R] may not be perceived as good a match to Spanish /r/ as German [r] (see also, e.g., Guion et al., 2000; Lancaster and Gor, 2016; for effects of degree of phonetic match in cross-language perception). Lastly, German [ʁ] has several possible mappings if, in addition to the phonological/functional match, acoustics and articulation are considered. The phonetically closest Spanish phones to German [ʁ] are the Spanish velar approximant [ɣ] (the intervocalic allophone of /g/) and—as discussed above—the velar fricative /x/. **Table 1** summarizes all potential within- and cross-language mappings of L1 and L2 phones based on phonological vs. phonetic characteristics.

To address the potential relationship between cross-language perception and L2 phoneme identification, we tested whether the individual learners' cross-language perception pattern correlates with their performance in the identification of [ʁ] as /r/. If phonological mapping matters the most for L2 sound learning, then according to PAM-L2, the best performance in the phoneme identification task would be expected by those learners who have established an allophone [ʁ] within their Spanish-German phonological category /r/. The establishment of [ʁ] as an allophone of /r/ would manifest as a consistent mapping of German [ʁ] to Spanish /r/ in the cross-language perception task. However, PAM-L2 also predicts that L2 phones that are not straightforwardly mapped to a single L1 phoneme should be reasonably easy to learn if no other L2 category is associated with the same set of L1 categories (i.e., “no L1-L2 phonological assimilation”; Best and Tyler, 2007:30). This could also be the case for German [ʁ], and such an outcome would be reflected in the cross-language perception task by learners picking several Spanish phonemes that do not necessarily have to be Spanish /r/. In that case, better performance in phoneme identification might be expected to correlate with response consistency in the cross-language task irrespective of the specific L1 phonemes chosen.

To explore these predictions, two types of correlation analyses were conducted between correct responses in the identification task for German words starting with [ʁ] and measures from cross-language perception. The first measure from cross-language perception was the proportion of (phonological) mappings of German [ʁ] to Spanish /r/. The other measure was a more general measure of mapping consistency that takes into account the overall distribution of answers across response options: Entropy (Shannon, 1948). Entropy provides a measure of consistency/agreement that has been used in Linguistics, for instance, to test listeners' consistency in projecting the end of turns in a conversation depending on different prosodic conditions (De Ruiter et al., 2006). Entropy was used to assess how consistently participants matched a given German allophone of /r/ to one vs. more Spanish phonemes in the cross-language perception task. Importantly, as a global measure accounting for overall response patterns, entropy avoids the need for arbitrary decisions about cut-off points that have previously been used to classify a given L2 phone as an L1 phoneme (e.g., Harnsberger, 2001; Bundgaard-Nielsen et al., 2011; Tyler et al., 2014; Faris et al., 2016). For instance, Faris et al. (2016) used a criterion that, in a cross-language perception task, a non-native vowel had to be mapped onto a particular L1 vowel at least 50% of the time in order to be counted as consistently “categorized” as this vowel, while Tyler et al. (2014) used a 70% and Harnsberger (2001) even a 90% criterion. Entropy was therefore introduced as a general measure of classification consistency in order to avoid having to resort to a pre-established decision criterion.

In sum, the aim of the present study was to assess how Spanish learners of German deal with free allophonic variation in L2 phoneme perception, that is, in a scenario in which phoneme identity cannot be predicted based on one specific set of acoustic cues or phonotactic information. Specifically, we investigated whether German [r], [R], and [ʁ] would differ in how well they are identified as the phoneme /r/ depending on their

phonetic similarity to the corresponding L1 Spanish phoneme /r/ and to another, functionally distinct German phoneme: /h/. We also tested the learners' cross-language mapping patterns for the three German allophones in question and asked whether identification performance relates to how learners match the German allophones to one vs. more native-language phonemes.

METHODS

Participants

Twenty-two native speakers of Castilian Spanish (13 female, 9 male) participated for a small monetary compensation (9 Euro). They all gave written informed consent in accordance with the Declaration of Helsinki. The study was carried out in accordance with the recommendations of the University of Munich and the funding agency, which required no separate ethics approval. Participants were recruited via advertisements on notice boards at the University and via social media with the requirements of having grown up in Spain, being between 18 and 40 years of age, having learned German as an adult and speaking German at medium to high proficiency. According to a questionnaire that participants filled out after completing the experiments, participants in the recruited sample were between 19 and 38 years old (average 27.2, $sd = 5.1$) and had been living in Germany between 0.25 and 5.5 years (average 2.7, $sd = 1.9$). They had started learning German as adults between 15 and 33 years of age (average 24.2, $sd = 4.9$). Sixteen of the twenty-two learners reported using German relatively frequently on a daily basis (mean rating of 2, $sd = 0.8$ on a scale from 1 = very frequently to 7 = very little); only six learners reported using German less frequently (4 to 6 on the 7-point scale). Eight participants had learned Galician (3) or Catalan (5) as a second L1. Since neither of these languages contains additional sounds similar to the critical sounds of the study, this was considered acceptable. However, data from two participants (one female, one male, one among the frequent users of German) were excluded because one of them had learned Basque and the other Arabic as a second L1. Both languages have the phoneme /h/ (Al-Ani, 1970; Hualde, 1991), hence including their data may have distorted the results.

All participants reported to have had several years of instruction in English. Therefore, they were likely to have had some exposure to /h/. However, unlike in the case of the two excluded participants, it was not expected that learners had mastered /h/ at a native-like level. About half the participants reported to have also had some instruction in French though none reported being able to speak this language. Since French is known for its realization of /r/ as a uvular fricative [ʁ] or [χ] (Gendrot et al., 2015), subsets of data were analyzed to assess the impact of learners' contact with French. Since separate analyses for participants with and without instruction in French rendered overall similar patterns of results, only the pooled data will be reported.

Material

Six German high-frequency word pairs that minimally differed in their initial phonemes /h/ or /r/ were selected such that Spanish learners of German would likely be familiar with them. Three Spanish learners of German who later did not participate in the

main experiments but had a similar background and learning history as the other participants confirmed their familiarity with these words. The word pairs were *Hose-Rose* ("pants"- "rose," log-frequencies per million of 3.01 and 3.06, respectively, as assessed in the SUBTLEX-DE corpus, Brysbaert et al., 2011), *Hund-rund* ("dog"- "round," log-frequencies of 3.54 and 2.59), *heiß-Reis* ("hot"- "rice," log-frequencies of 3.45 and 2.66), *heißen-reißen* ("to be called"- "to tear," log-frequencies of 3.45 and 2.58), *Hand-Rand* ("hand"- "edge," log-frequencies of 3.78 and 2.43) *Haus-raus* ("house"- "out," log-frequencies of 4.01 and 4.25).

Nine female native speakers of German, aged 23 to 30 (mean 25.5, $sd = 2$) were selected as speakers. They all grew up and lived in the South of Germany. None had learned any other language before the age of 8. The speakers were selected to differ in their natural use of German allophones of /r/. Five speakers were selected to mainly use the uvular fricative [ʁ] that is considered the standard and most frequent German variant (Ulbrich, 1973; Wiese, 2003; Ulbrich and Ulbrich, 2007). Two additional speakers were selected to produce a uvular trill [R] and two an alveolar trill [r] representing additional variants that are sometimes encountered in the South of Germany where the study took place. Given the prevalence of [ʁ], more speakers were recorded using this variant than each of the other variants. All words were recorded in a sound attenuated booth using a diaphragm microphone (Neumann Microphone, type TLM 103) and SpeechRecorder software (Draxler and Jänsch, 2004), which stored each utterance as a separate wav file onto the hard drive of a desktop computer. Speakers were seated ~30–40 cm in front of the microphone and read the words from a computer monitor. Production data were sampled at 44.1 kHz with 16-bit quantization. Each word was spoken in isolation four times per speaker in randomized order and the best two tokens per word per speaker were selected according to recording quality, duration (i.e., speech rate) and clarity of the intended /r/ variant as assessed by a phonetically trained native speaker of German. Selected tokens were normalized in root-mean-square amplitude using Praat (Boersma and Weenink, 2017).

In order to ensure that the stimuli sounded natural and were fully intelligible, 7 native speakers of German were asked to perform the identification task that is described in the following section. German speakers were able to correctly identify all three allophones of /r/, as well as /h/, with >99% accuracy ([r] and [R] = 100%, [ʁ] = 99.2, and /h/ = 99.1), indicating that the stimuli were appropriate and not perceptually confusable for native speakers of the language.

Procedure

Spanish participants completed all tasks in the same order. They first performed the cross-language perception task, followed by the minimal-pair identification task. The cross-language perception was conducted before identification because the identification task focuses learners' attention on the phonological function of the phones. A focus on phonological properties could in turn have affected choices in the cross-language perception task such that listeners preferred matching stimuli by phonological rather than phonetic similarities. By asking them to do the cross-language perception task first we reduced this bias as much as possible. However, given that the main task of interest

was the minimal pair identification, this task will be reported first in the procedure and results section. After the perception tasks, participants were asked to fill out a language background questionnaire. The complete session lasted ~ 1 h.

Identification Task

Participants received written instructions in Spanish that asked them to identify words of German minimal pairs. On each trial, the two German words of a minimal pair were presented orthographically on the left and right side on the screen. According to German orthography nouns were spelled with an initial capital letter. Over the course of the experiment the side on which words starting with /h/ and /r/ were presented was stable for one participant but counterbalanced across participants. Seven hundred milliseconds after the response options appeared on the screen, one word of this minimal pair was presented auditorily over headphones (Beyerdynamic, DT770) at a comfortable listening level. Participants were asked to press the 0-key on the computer keyboard if they thought the word they heard corresponded to the right word on the screen, and the 1-key if they heard the word on the left. They were not specifically encouraged to provide their responses as fast as possible. No feedback was given. Five hundred milliseconds after the button press the next trial started automatically. This timing was assessed through piloting in order to be perceived as comfortable. All words and voices were presented in a different randomized order to each participant. Participants completed a total of 432 trials (6 minimal pairs × 2 tokens per word × 9 speakers × 2 repetitions). The experiment was implemented in PsychoPy2 (version 1.83.01; Peirce, 2007) and took participants approximately 15 min to complete.

Cross-Language Perception Task

In the cross-language perception task participants were asked to match the initial phones of the German minimal word pairs to a set of Spanish phonemes that were chosen to be acoustically and articulatorily sufficiently similar to German /h/ and the different variants of /r/ so as to serve as potential phonetic targets for perceptual mapping. We decided to use existing German words rather than non-words, firstly, to keep the sound material identical to the identification task, and secondly, because with non-words it would not have been possible to disentangle phonetic vs. phonological influences in L1-L2 mappings, as non-words lack a lexical, hence phonological representation. Note also that participants were learners of German who could rely on lexical knowledge when using their L2 (unlike naïve listeners that are often used in cross-language perception studies as cited, e.g., in Best and Tyler, 2007). The Spanish phones that listeners were asked to match the German phones to were /t/, /r/, /k/, /x/, and [ɣ] as represented in word-medial position in the words *paro*, *parra*, *paco*, *paja*, *pago* (“unemployment,” “grapevine,” a common Spanish proper name, “hay,” “payment”), respectively. Note that all phones are phonemes in Spanish except for [ɣ], which is an intervocalic variant of the phoneme /g/. Since native speakers of Castilian Spanish produce this variant form very consistently in intervocalic position (Carrasco et al., 2012) it was assumed

that when choosing *pago* participants were indeed linking their response to the variant form [ɣ].

Participants received written instructions in Spanish and could ask the experimenter if anything was unclear. They were presented the words displayed in their orthographic form on a computer screen. They were asked to pick the Spanish word that contained a “sound” that sounded most similar to the “initial sound” of the German word. They were made aware of the fact that critical phones in German were in word-initial position but in word-medial position in Spanish. The letters corresponding to critical phones in the Spanish words were highlighted. The words were distributed randomly over six fields on a computer screen (with one field remaining empty) such that positions were consistent for one participant but varied between participants. This was as to avoid effects of position in the choice of response options.

On each trial, one second after the appearance of the words on the screen, one of the German words was presented auditorily over headphones (Beyerdynamic, DT770) at a comfortable listening level. Participants’ task was to carefully listen to the word—specifically the word’s initial phoneme—and indicate by mouse click which of the words presented on the screen contains a phone similar to the one in the audio stimulus. Five hundred milliseconds after participants had chosen a Spanish word, this word was presented in the middle of the screen with the question *¿Cuánto se parece al sonido que has elegido?* (“How similar is this to the sound you chose?”) above and a 5-point rating scale below. Simultaneously with the appearance of the rating scale, participants heard the German word once again. Now their task was to decide how well the phone in the Spanish word that they had picked matched the initial phone of the German word. One was labeled as *mucho* (“very”), 5 as *muy poco* (“very little”; note that for analyses we re-coded these values according to the more common higher-is-better scheme). The answer was logged by clicking with the mouse onto one of the five points on the scale. Upon the mouse click the next trial started automatically after an inter-trial interval of 200 ms. The timing was again established through piloting so as to be perceived as comfortable. Response speed was not emphasized. Participants completed a total of 216 fully randomized trials (6 minimal pairs × 2 tokens per word × 9 speakers). The experiment was again run in PsychoPy2 (version 1.83.01; Peirce, 2007) and took approximately 25 min to complete.

RESULTS

Identification

Figure 1 shows the results for the identification of the German minimal pairs split by /h/ and the three allophones of /r/. As can be seen from the scale on the y-axis (starting at 80%), participants were very accurate overall. Identification of words spoken with the two trilled allophones of /r/ was close to ceiling (both > 99% correct), words with /h/ were identified with about 96% accuracy and the identification rate of words with the fricative [ɣ] was the lowest with still about 92% correct. The pattern of reaction times for the different allophones of /r/ and /h/ mirrored the accuracy pattern, suggesting that accuracy data were not due to

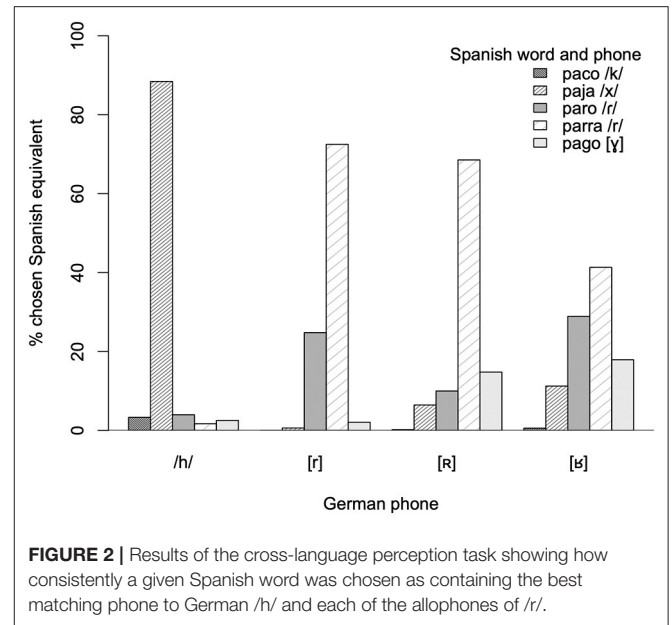
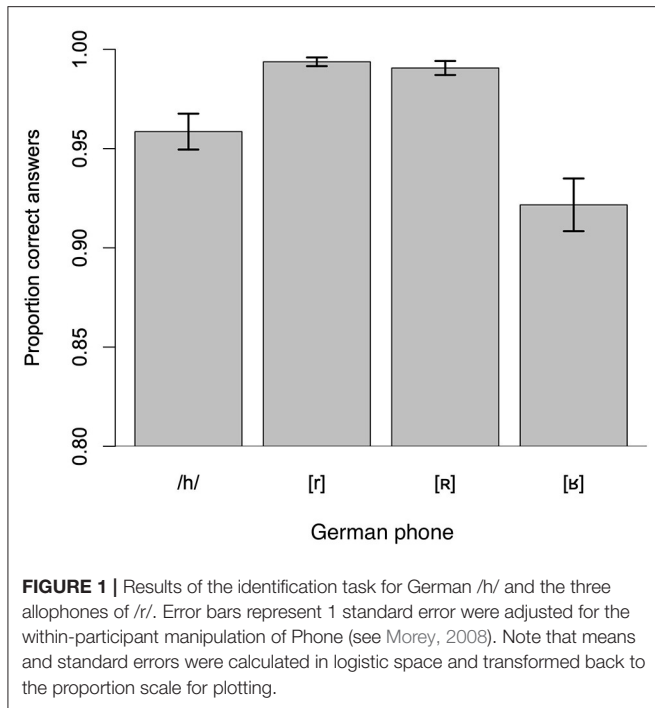


TABLE 2 | Results of the general linear hypothesis testing for the generalized mixed-effects model for correct responses in the minimal pair identification task.

Hypothesis == 0	B	z	p
[R]-[r]	-0.02	-0.028	0.99
[ɐ]-[r]	-2.56	-4.62	<0.001
/h/-[r]	-2.07	-2.68	<0.05
[ɐ]-[R]	-2.54	-4.30	<0.001
/h/-[R]	-2.04	-2.48	0.06
/h/-[ɐ]	0.49	0.8	0.85

Adjusted p-values are reported.

a speed-accuracy trade-off: When reaction time was measured from acoustic target onset, the two trilled allophones of /r/ were responded to the fastest, in 792 ms for [r] and 807 ms for [R]; /h/ was responded to on average 851 ms after target onset and the reaction to [ɐ] was slowest with a response time of 962 ms (counting correct trials only). However, since participants were not explicitly encouraged to respond fast, only accuracy data were analyzed further.

To test whether differences in accuracy between the different phones were statistically significant, a linear-mixed effects model was fit with a logistic linking function using the lme4 package (version 1.1-14; Bates et al., 2014) in R (version 3.4.2; R Core Team, 2017). Accuracy was the dependent variable and Phone the fixed factor with the four levels /h/, [r], [R], and [ɐ]. Participant and word/target were included as random factors with a random slope for Phone over participants. This model appeared to be a better fit relative to models with less complex random-effects structures as assessed via log-likelihood ratio tests

using the anova() function. To assess differences in identification performance between the different realizations of /r/ and /h/, we subjected the model to general linear hypothesis testing using the glht() function from the multcomp package (Hothorn et al., 2016) in R, using the Tukey Method to account for multiple comparisons. **Table 2** shows the results for the pairwise comparisons. Words that were produced with [ɐ] were identified significantly worse than words produced with all other sounds. Moreover, words produced with /h/ were identified significantly worse than words with [r].

In order to exclude the possibility that the worse identification of words starting with [ɐ] was due to the specific speakers that we chose (note that in our design the allophones of /r/ were speaker-specific), it was additionally tested whether differences could be found for the identification of the different speakers' productions of /h/-initial words. If participants' difficulty in identifying [ɐ] depended on the speakers more generally, rather than on their specific production of /r/, then one could speculate that similar difficulties would arise for these speakers' productions of /h/ relative to the other speakers' /h/. To test this possibility the same linear mixed-effects model as described above was fit with only the subset of data in which /h/ was the intended target. Feeding this model again to linear hypothesis testing revealed that speakers producing [ɐ] did not differ from the other groups in how well their /h/ productions were identified ([ɐ]-[r]: $b = -0.58, z = -1.74, p_{(adjusted)} = 0.19$; [R]-[ɐ]: $b = -0.4, z = -1.73, p_{(adjusted)} = 0.19$). Only the /h/-initial words by speakers using a uvular trill variant of /r/ were identified significantly worse than by speakers using an alveolar trill ($b = -0.98, z = -2.8, p_{(adjusted)} < 0.05$). While this latter result is hard to explain, the main result here is that learners' difficulties with identifying [ɐ]-initial German words are not due to these speakers' poor intelligibility more generally.

Cross-Language Perception

Figure 2 shows the cross-language perception patterns for German /h/ and the three allophones of /r/ to the Spanish phones /k/, /x/, /r/, /r/ and [ɾ] as represented in the words *paco*, *paja*, *paro*, *parra*, *pago*. As can be seen from the figure, German /h/ was mainly categorized as Spanish /x/-as expected—however, the goodness of fit was judged as moderate-to-fair with an average of 3.47 on the 5-point scale (5 = best).

All three German allophones of /r/ were most frequently categorized as Spanish /r/ compared to the other response options (see **Figure 2**). That is, all three allophones were most frequently mapped onto the functionally/phonologically matching category. However, differences were found between the German allophones of /r/ in terms of how frequently Spanish /r/ was chosen. These differences were analyzed using a generalized linear mixed-effects model with /r/-response as the dependent variable (coded 1 if /r/ was the chosen response option, else coded 0), Allophone as fixed factor with three levels [r], [R], [ɾ], random intercepts for participant and word, and a random slope for Allophone over participants. Comparisons between allophones were again assessed by general linear hypothesis testing, using Tukey method to account for multiple comparisons. Results showed that the Spanish phoneme /r/ was chosen significantly less often for the German fricative [ɾ] than either of the two trilled allophones ([ɾ]-[r]: $b = -1.88$, $z = -5.55$, $p < 0.001$; [ɾ]-[R]: $b = -1.63$, $z = -6.02$, $p < 0.001$). Responses for the two trills did not differ from one another ([r]-[R]: $b = 0.25$, $z = 0.68$, $p = 0.775$). A linear-mixed effects model with the same structure was run with goodness ratings for the choice of Spanish /r/ as dependent variable and a Gaussian linking function to account for the quasi-continuous variable. Results showed similar patterns as the analysis above: the German uvular fricative [ɾ] (mean rating = 2.98) showed significantly lower ratings for goodness of match to Spanish /r/ than the two trilled variants [mean rating [r] = 3.88, comparison [ɾ]-[r]: $b = -0.86$, $z = -9.38$, $p < 0.001$; mean rating [R] = 3.33, comparison [ɾ]-[R]: $b = -0.39$, $z = -3.11$, $p = 0.005$]. Unlike for the overall proportion of /r/-responses mapping the German trills onto Spanish /r/, the two trills significantly differed in their ratings for goodness of match ([r]- [R]: $b = 0.47$, $z = 2.98$, $p = 0.007$).

In addition to the proportion /r/-responses and goodness ratings, differences were found in the cross-language perception of the German allophones of /r/ in terms of the overall pattern of response options. That is, the German allophones of /r/ differed in how consistently participants picked one single Spanish phone as the best match as compared to many phones. As can be seen in **Figure 2**, German [r] was mapped to Spanish /r/ in 72% of the cases with just about 25% of tap responses. [R] was mapped to /r/ in 69% of the cases followed by [ɾ] and /r/ with 15% and 10% of the choices, respectively. Finally, [ɾ] was mapped to /r/ in only 41% of the cases with /r/ and [ɾ] being chosen in 29 and 18% of the cases. Importantly, large individual differences were found in the preferred mapping patterns for [ɾ] as is illustrated in **Table 3**.

In order to capture the distribution and consistency of participants' answers for each allophone of German /r/ and the

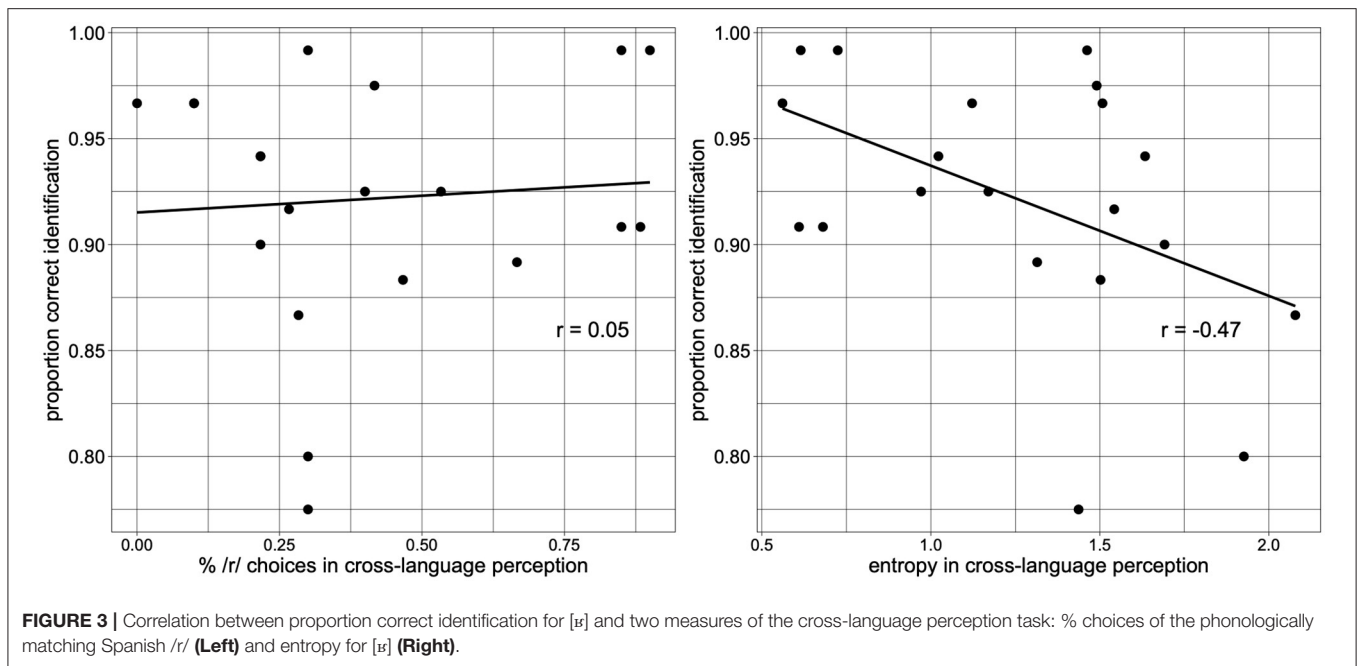
TABLE 3 | Percent chosen option for cross-language perception of the German fricative variant [ɾ] per participant and each participant's measure of entropy as a measure of consistency of choice.

Participant	Percent chosen response option for [ɾ]				Entropy
	Paco/k/	Paja/x/	Paro/r/	Parra/r/	
pp_4		90	3	7	0.56
pp_8		10	3	76	1.12
pp_11		3	33	53	1.51
pp_12		17	53	8	1.69
pp_20	5	2	58	13	1.63
pp_1		5	73		1.2
pp_10			45	28	1.54
pp_5	2	20	25	25	2.1
pp_7		8	3	58	1.44
pp_19		15	20	35	1.93
pp_16			52	18	1.46
pp_2			60		0.97
pp_6	2	7	48	2	1.49
pp_15		12	40	2	1.50
pp_17			43	3	1.17
pp_13		12	12	20	1.31
pp_18		3	12		0.72
pp_9		15			0.61
pp_3		7	2	3	0.68
pp_14	3		2	5	0.61

Data are sorted in ascending order for the percent chosen /r/ (i.e., parra). Empty fields indicate that this option was never chosen. pp_00 indicates the participant number.

phoneme /h/ we calculated a measure of entropy (Shannon, 1948) using the R package 'entropy' (Hausser and Strimmer, 2009). Entropy provides a measure of consistency/agreement that takes into account not only how often one particular response option was chosen, but the distribution of answers across response options. The interpretation is that if participants had picked only one and always the same Spanish phone for a given German phone the entropy would be zero. In other words, the higher the value of entropy, the more distributed the answers were across the response options. Participants' mean entropy for /h/ was 0.56, for [r] 0.74, for [R] 0.98, and for [ɾ] it was 1.25 (see **Table 3**).

To assess differences in how consistent vs. distributed the cross-language mapping patterns were for the different allophones of German /r/, pairwise comparisons were calculated as for the previous measures. An lme-model (Pinheiro and Bates, 2000) was fit on the entropy data per participant with Allophone ([r], [R], [ɾ]) as fixed factor and accounting for the fact that Allophone was manipulated within participants. This model was then subjected to general linear hypothesis testing, using the Tukey method to adjust for multiple comparisons. Results showed that the entropy for the fricative variant [ɾ] was significantly higher, that is, the response pattern was more distributed across response options for [ɾ] than for the two trilled allophones of /r/ ([ɾ]-[r]: $b = 0.52$, $z = 5.14$, $p < 0.001$; [ɾ]-[R]: $b = 0.36$, $z = -3.60$, $p < 0.001$), but the two trills did not differ from one another ([r]- [R]: $b = -0.15$, $z = 1.5$, $p = 0.271$).



Correlation Between Tasks

To assess whether the degree of phonological relative to phonetic mapping of L2 allophones onto the L1 influences how well these allophones are identified as a given phoneme in the L2, we tested whether the cross-language perception pattern would correlate with individual learners' performance in the identification of [ʁ] as /r/. The focus was on [ʁ] because this allophone shows the smallest phonetic difference to another German phoneme (i.e., /h/) and the largest phonetic difference to the phonologically matching Spanish phoneme /r/. In the experiments, [ʁ] was the least well-identified allophone of /r/ and showed the most variation between participants in terms of cross-language perception patterns (see Table 3).

Two variables from the cross-language perception task were considered in relation to L2 phoneme identification. First, it was assessed whether the proportion of choosing the functionally/phonologically matching Spanish category (i.e., /r/) in the cross-language perception task when hearing the [ʁ] variant correlated with the percent correct identification of German words starting with [ʁ]. However, a significant correlation could not be found ($r = 0.07$; $p = 0.756$). Second, it was assessed whether more consistent cross-language mapping of [ʁ] to a single L1 phoneme over more than one response option would correlate with better identification of L2 words (i.e., entropy). Here a significant medium-to-large correlation (Cohen, 1988) was found ($r = -0.47$; $p < 0.05$): the higher the entropy score for a given participant, that is, the less consistently [ʁ] was mapped to a single Spanish phoneme, the less accurate this participant was in correctly identifying words starting with [ʁ]. Figure 3 shows the correlation for the two measures, proportion /r/-responses in the left panel and entropy in the right panel⁴.

⁴Two participants had accuracy rates in the identification of [ʁ] around or below 80%. Given that the scale of the plots was adjusted so that it captured the relatively

In addition to testing for correlations between tasks, we asked whether self-reported measures of proficiency as assessed in the language-background questionnaire could predict performance in L2 phoneme identification and/or cross-language perception. We examined several combinations of factors related to L2 language use and L2 proficiency, for instance, an index combining five values of self-estimated experience and proficiency (self-reported frequency and skills in listening and speaking, as well as the learners' self-estimated accent) that had previously been shown to predict learners' ability to rate the accent of peer learners in their L2 (Eger and Reinisch, 2019, for German learners of English). However, this index did not correlate with the identification accuracy of German words produced with [ʁ] ($r = -0.18$; $p = 0.459$) or cross-language perception (i.e., entropy: $r = -0.044$; $p = 0.859$) and neither did any other measure from the questionnaire.

DISCUSSION

The present study set out to test how Spanish L2 learners of German deal in perception with allophonic variant forms that are not specific to a given word position but vary freely with some constraints across speakers and regional varieties. Using the case of the German phoneme /r/, which can variably be produced as [r], [R], or [ʁ], we asked to what extent phonetic factors influence L2 perception of an L2 phonological category that is common to L1 and L2. This was achieved by comparing the identification of the different allophones of German /r/ relative to the German phoneme /h/, which is phonetically similar to [ʁ] but much less

small differences in accuracy observed, rates for these two participants might seem to be outliers. However, as both values fall within 2.5 standard deviations from the mean of all participants, they were not considered outliers and were kept in all analyses.

so to [R] or [r]. If phonetic factors played a role in L2 perception in addition to L2 phonology, /h/ and [ʁ] could be predicted to be confusable⁵ for learners.

Results of the identification task, however, showed that Spanish learners of German were overall very accurate at identifying the different German allophones of /r/ as the intended phoneme (>90% correct). Confusion was hence small in general. Nevertheless, differences were found in how well each of the three allophones of /r/ was identified. Words produced with [ʁ] were identified correctly significantly less often than words with [R] and [r]. Results of the identification task hence suggest a major influence of the phonological category in L2 identification (cf. >90% correct) with some additional role for the phonetic match of a given variant form to the associated phonological category vs. other L2 phonemes. A potential caveat here is that the larger number of speakers producing the [ʁ] allophone than [r] or [R] could have caused greater variability and hence increased difficulty for [ʁ]. However, whilst we cannot rule out such an influence, we consider it unlikely that this can fully account for our results. Learners' accuracy rates in identifying [ʁ] as /r/ across speakers were very similar (i.e., 91.3, 91.9, 90.8, 91.7, and 95.2%, respectively) and German participants in the pilot experiment were at ceiling.

The second angle from which the role of phonetic match in L2 perception was addressed concerned the perception of the German allophones of /r/ and /h/ relative to the learners' L1 Spanish, as assessed in a cross-language perception task. It was assumed that even for learners of medium-to-high proficiency, the L1 likely plays a role in L2 perception. Spanish has a phoneme /r/ which can be considered a phonological and functional match to German /r/. However, since Spanish /r/ is typically produced as an alveolar trill, the German allophones differ in their phonetic match to Spanish /r/. If learners had only relied on the phonological function of the German allophones in this task, then the expected choice would have been the phonologically matching Spanish /r/. Indeed, Spanish /r/ was the most frequent choice for all allophones of German /r/. However, substantial differences were found between allophones in the proportion of Spanish /r/ responses and their goodness ratings. German [r] was most frequently mapped to Spanish /r/ and rated the best fit, followed by [R], which was followed by [ʁ]. In line with the results of the identification task discussed above, this finding outlines again a relevant role of phonetic influences on L2 phoneme perception.

Considering the distribution of chosen responses for the German allophones in cross-language perception, a number of additional findings deserves discussion. First, German [r], which had been recognized as an allophone of /r/ in the identification task with >99% accuracy, was not exclusively mapped to Spanish /r/ in cross-language perception but also to /r/ in about 25% of

the cases. This suggests that (at least some) listeners appeared to attend to fine phonetic detail that typically differentiates the German from the Spanish alveolar trill: the typical number of occlusions which is 1–3 for German (Wiese, 1996) but 3–5 for Spanish (Navarro-Tomás, 1916; Blecua, 2001; Hualde et al., 2010), and the associated difference in duration (shorter for German). This difference in fine phonetic detail then may explain why the German alveolar trill is sometimes perceived as similar to the Spanish tap.

Second, in cross-language perception of the German uvular trill [R], listeners responded most frequently with the phonologically matching option /r/, with which it shares the trilled manner of articulation, but this was not the only response: at least some of the time participants mapped German [R] to Spanish [ʁ] and /x/. Notably, these two phones share the back place of articulation with [R]. This descriptive pattern suggests that for L2 learners (in addition to the phonological match) the acoustic characteristics reflecting manner of articulation appear most important in their decisions. This is likely because the acoustic correlates for manner (here the trills) are perceptually much more salient than acoustic cues to place (see Martin and Peperkamp, 2017, for other aspects of manner). These results are in line with a study by Colantoni and Steele (2007) who show that in L2 production, English learners of French tend to acquire those aspects of a segment's articulation first that are perceptually salient.

Finally, the German allophone [ʁ] was the most variable allophone in terms of response patterns in the cross-language perception task, within as well as between participants (see **Table 3**). As mentioned in the introduction, we calculated a measure of entropy in order to quantify the consistency with which participants picked one or more options for this allophone in cross-language perception. Interestingly, entropy was the only measure of the cross-language perception task that correlated with learners' accuracy in the identification task: the more consistently participants mapped [ʁ] in cross-language perception to a single Spanish phoneme, the better they were at identifying [ʁ] as /r/ in the identification task. Importantly, the phoneme that was consistently selected did not always have to be the phonologically matching Spanish /r/, since the mere proportion of selecting Spanish /r/ for the [ʁ] variant did not correlate with the identification results for that allophone.

An issue regarding the cross-language perception task is that the metalinguistic affordance of this task might reflect not only participants' decisions based on language processing, but also the outcome of individual strategic behavior. That is, that there may be differences across participants in whether they report L1-L2 associations based on a phonetic level of processing (i.e., favoring non-/r/ responses for German [ʁ]) or a more abstract phonological level (i.e., favoring /r/ responses for German [ʁ]). It could be argued that strategic behavior may have been additionally fostered by our choice to use real German words rather than non-words. In spite of these limitations, it must be noted that cross-language perception tasks have frequently been used to evaluate L2 perception and the predictions of L2 learning models (see e.g., studies reported in Best and Tyler, 2007). However, even though task-related effects are a likely

⁵Note that by testing identification of the German contrast /r/-/h/ the greatest confusion was expected for the allophone [ʁ] (rather than [r] and [R]) by design. However, testing additional contrasts, for instance, German /r/ vs. /d/ or /g/ would not have led to similar expectations about confusions between the stops and [r] or [R], despite their respective match in place of articulation. This is because the trilled manner of articulation of [r] and [R] is perceptually quite salient and, importantly, /d/ and /g/ are also phonemes in the learners' L1 Spanish.

possibility, strategic behavior does not easily explain the present results regarding the correlation of cross-language mappings of [ɾ] and performance with this allophone in the identification task. If consistency in mapping as measured by entropy was purely determined by learners' strategic behavior, one could wonder why entropy would correlate with better performance in the identification task.

What do the findings on the relation between the consistency in cross-language perception and identification mean with regard to models of L2 sound learning and the acquisition of an L2 phonological inventory in general? PAM-L2, which we used to base our specific predictions on, makes two assumptions that are potentially relevant to the perception of the German allophone [ɾ] for our Spanish learners, whose native realization of /r/ is an alveolar trill. First, as discussed in the introduction, the model suggests that if a phonological category exists in the L1 and L2 but the typical realization differs, then a separate phonetic category within the common L1-L2 phonological category can be established (see the example of English /ɹ/ vs. French /ʁ/ in Best and Tyler, 2007). However, the present case for the allophone [ɾ] was somewhat more complex than the mere mismatch in phonetic form between two functionally equivalent phonological categories in the L1 and L2, as German has two additional allophones that both match the phonetic form of the typical realization of Spanish /r/ to a greater extent (and in addition [ɾ] is phonetically closer to another L2 phoneme /h/ which was the contrast tested in the present study). Considering that learners correctly identified [ɾ] less frequently than the two trilled variants of /r/, our results suggest that, at least in the case of multiple variant forms of the same phoneme, the lack of a robust phonetic match between an L2 allophone and the functionally matching L1 phoneme has consequences for that allophone in speech perception in spite of the phonological match existing between L1 and L2 phones.

Second, PAM-L2 suggests that, if a given L2 phone is not straightforwardly mapped onto any given L1 phonological category and no other L2 phone shows the same pattern of (non-)mapping to the same L1 phonemes, then the L2 phone should be learned well. This could again apply to [ɾ], the allophone of German /r/ that had the most variable response pattern in the cross-language perception task. In this respect, what our results show is that, perhaps surprisingly, the strength of the cross-language mapping between [ɾ] and only the phonologically matching L1 /r/ did not predict participants' identification accuracy for this allophone. Instead, what did predict identification accuracy was the overall consistency in cross-language mapping as measured by entropy, a measure that took into account not only /r/ responses but all responses provided. The more consistently responses were associated with only one Spanish phoneme, be it /r/ or another response option (see **Table 3**), the more accurate the identification of [ɾ] was. Importantly, this outcome contradicts the initial prediction that spread out patterns of responses for [ɾ] (i.e., "no L1-L2 phonological assimilation" scenario) would relate to better identification. However, note that, from the present results, it cannot be ruled out that cross-language perception patterns for other untested L2 phones may overlap with those for [ɾ], even

though that seems unlikely given the overall strong phonological component in the meta-linguistic cross-language perception task (see e.g., the assimilation pattern of /h/, which is phonetically similar to [ɾ]). In any case, the critical finding here is that entropy predicts identification accuracy, which has the potential to guide future methodological choices attempting to capture cross-language mapping consistency without the need for arbitrary criteria as to how often an L2 phone has to be categorized as a given L1 phoneme in order to count as "consistently mapped/assimilated" to this phoneme (e.g., Harnsberger, 2001; Bundgaard-Nielsen et al., 2011; Tyler et al., 2014; Faris et al., 2016).

Looking at L2 learning in general, thereby capitalizing on the identification task, we find strong phonological effects in L2 learning that to some extent are modulated by phonetic factors. In the present study we tested Spanish learners of German who had started learning German as adults. They were all living in Southern Germany, where they were likely exposed to all allophones of German /r/ under investigation. However, since at the time of participating they covered a reasonably wide range of proficiencies according to self-report, it could have been expected that participants who reported more contact to German in everyday life performed better in recognizing variant forms than those who reported less contact, but this is not what was found. Self-reported measures from the language-background questionnaire were not correlated to performance in identification or cross-language perception. This could be because self-reported measures are more strongly related to general indicators of overall proficiency than to performance in phonetic/phonological processing. Specifically, previous studies have shown that self-reported proficiency relates to performance in tasks that have a strong focus on lexical processing—in perception as well as production (e.g., Llompарт and Reinisch, 2019a,b)—rather than to the perception of sound contrasts in tasks with a very clear phone-level focus, as in the present study. This is most likely related to the fact that learners who are relatively poor at distinguishing between L2 phones when focusing on higher-level processing may still perform very well in an identification task that focuses their attention on the critical phones (e.g., Sebastián-Gallés et al., 2005; Díaz et al., 2012; Darcy et al., 2013; Llompарт and Reinisch, 2019a,b). Despite the present learners' variability in their language-learning history, the lack of relation between self-estimated proficiency and performance in the present study may hence not be so surprising.

A question that remains to be answered is the role of allophonic variant forms in the process of establishing an L2 phonological category /r/ that encompasses all three allophones: those similar to the learners' L1 as well as the new sub-category [ɾ]. While our study is unable to speak to the learning process itself, we can use our results to speculate. One possibility is that the existence of L2 allophones that are phonetically close to the L1 (here the L2 German allophones [r] and [R] that are close to the typical realization of L1 Spanish /r/) may to some extent hinder the association of the additional, phonetically distinct allophone [ɾ] with the intended phoneme /r/. This could in part be driven by the fact that L2 phonological categories including more than one allophone have to be acquired in perception and production

but, in the case of free allophonic variation, the affordances are rather different in the two modalities. While in production the transfer of the Spanish alveolar trill to German may be sufficient, in perception all allophones have to be recognized.

Alternatively, the existence of L2 allophones matching the L1 category (i.e., German [r] matching the typical realization of Spanish /r/) may help strengthen the phonological link between German /r/ and Spanish /r/ on the grounds of phonetic similarity. That is, in comparison to a scenario in which learners have to establish a phonological link between a non-native category /r/ that is always realized as [ʁ] and a native category /r/ that is always realized as [r], having [r] as a commonly heard allophone could be a useful tool to reinforce the phonological connection between the L1 and L2 rhotics. As a consequence, this may facilitate the incorporation of [ʁ] as an additional phonetic sub-category within the L2 phonological category /r/. This should eventually result in learners identifying the [ʁ]-allophone quite accurately in spite of the large phonetic dissimilarities, as it appears to be the case in this study.

However, in addition to the potential influence of the existence of common L1-L2 allophones, there is another factor that may have contributed to learners being overall quite good at identifying the phonetically distant [ʁ] variant: canonicity. [ʁ] has become the German standard pronunciation that is most frequently heard on national German media and is also indicated as the standard in language classes (as confirmed by our participants). Therefore, given that forms including canonical variants have been shown to have a special status in L1 listening (Deelman and Connine, 2001; Sumner and Samuel, 2005; Ranbom and Connine, 2007; Pitt et al., 2011; Sumner et al., 2014; Sumner, 2015; Llompарт and Simonet, 2018), one could expect something similar to occur in the L2, where (audio-visual) learning material typically covers the standard, more-canonical variety (see e.g., Baker and Smith, 2010, for a critical discussion of this issue). Further research could shed light on this issue by comparing the perceptual identification of potentially difficult L2 phones with and without the coexistence of other, “easier” allophonic variants.

In sum, the present study tested the perception of three allophones of German /r/ that can appear in free variation by Spanish learners of German. Results showed that identification of the allophones was crucially modulated by the phonetic similarities and dissimilarities between the allophones and i) the

most common L1 realization of the phonologically-matching L1 rhotic, and ii) the phonetic distance between the allophones in question and other L2 categories (i.e., /h/). In addition, correct allophone identification was found to relate to how consistently a given L2 variant was mapped to a single vs. many L1 phonological categories, with more consistency being linked to higher identification accuracy. All in all, these findings support the idea that both phonological *and* phonetic relations between the learners’ L1 and L2 are of importance in L2 phonological category identification in speech perception.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

ER and KJ designed the study. KJ conducted the study with technical support from ML. ER analyzed the data. ER and ML wrote the paper. All authors contributed to the article and approved of the final version. This research was conducted while all authors were at the Institute of Phonetics and Speech Processing at LMU Munich.

FUNDING

This project was funded by a grant from the German Research Foundation (DFG; grant nr. RE 3047/1-1) to ER. KJ received support from the Lehre@LMU programme of the University of Munich “Forschung entdecken: Förderung von Forschungsorientierung in der Lehre” (discovering research: support of research-oriented teaching; reference number: 870058-5_354_24.03.2017_Reinisch, Juhl).

REFERENCES

- Al-Ani, S. H. (1970). *Arabic Phonology: An Acoustical and Physiological Investigation*. The Hague, Netherlands: Mouton. doi: 10.1515/9783110878769
- Baker, W., and Smith, L. C. (2010). The impact of L2 dialect on learning French vowels: native English speakers learning Quebecois and European French. *Can. Mod. Lang. Rev.* 66, 711–738. doi: 10.3138/cmlr.66.5.711
- Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., et al. (2014). *Package ‘lme4’*. R Foundation for Statistical Computing, Vienna, 12.
- Best, C. T. (1994). “The emergence of native-language phonological influences in infants: a perceptual assimilation model,” in *The Development of Speech Perception: The Transition From Speech Sounds to Spoken Words*, eds. J. C. Goodman and H. C. Nusbaum (Cambridge, MA: MIT Press), 167–224.
- Best, C. T. (1995). “A direct realist perspective on cross-language speech perception,” in *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, ed. W. Strange (Timonium, MD: York Press), 171–204.
- Best, C. T., and Tyler, M. D. (2007). “Nonnative and second-language speech perception: commonalities and complementarities,” in *Language Experience in Second Language Speech Learning: In Honor of James Emil Flege*, eds. O.-S. Bohn and M. J. Munro (Amsterdam, NL: John Benjamins), 13–34. doi: 10.1075/llt.17.07bes
- Blecua, B. (2001). *Las vibrantes del español: Manifestaciones acústicas y procesos fonéticos* [Spanish rhotic consonants: acoustic characteristics and phonetic

- processes]. Unpublished (doctoral dissertation), Universidad Autónoma de Barcelona.
- Boersma, P., and Weenink, D. (2017). *Praat: Doing Phonetics by Computer*. Version 6.0.36, Available online at: <http://www.praat.org/>.
- Brysbart, M., Buchmeier, M., Conrad, M., Jacobs, A. M., Bolte, J., and Bohl, A. (2011). The word frequency effect: a review of recent developments and implications for the choice of frequency estimates in German. *Exp. Psychol.* 58, 412–424. doi: 10.1027/1618-3169/a000123
- Bundgaard-Nielsen, R. L., Best, C. T., and Tyler, M. D. (2011). Vocabulary size matters: the assimilation of second-language Australian english vowels to first-language Japanese vowel categories. *Appl. Psycholinguist.* 32, 51–67. doi: 10.1017/S0142716410000287
- Carrasco, P., Hualde, J. I., and Simonet, M. (2012). Dialectal differences in Spanish voiced obstruent allophony: costa rican versus iberian Spanish. *Phonetica* 69, 149–179. doi: 10.1159/000345199
- Cohen (1988). *Statistical Power Analysis for the Behavioral Sciences, 2nd Edn.* Hillsdale, NJ: Erlbaum.
- Colantoni, L., and Steele, J. (2007). Acquiring /[alveolar approximant]/ in context. *Stud. Second Lang. Acquisit.* 29, 381–406. doi: 10.1017/S0272263107070258
- Darcy, I., Daidone, D., and Kojima, C. (2013). Asymmetric lexical access and fuzzy lexical representations in second language learners. *Mental Lexicon* 8, 372–420. doi: 10.1075/ml.8.3.06dar
- De Ruiter, J. P., Mitterer, H., and Enfield, N. J. (2006). Projecting the end of a speaker's turn: a cognitive cornerstone of conversation. *Language* 82, 515–535. doi: 10.1353/lan.2006.0130
- Deelman, T., and Connine, C. M. (2001). Missing information in spoken word recognition: nonreleased stop consonants. *J. Exp. Psychol. Hum. Percept. Perform.* 27, 656–663. doi: 10.1037/0096-1523.27.3.656
- Díaz, B., Mitterer, H., Broersma, M., and Sebastián-Gallés, N. (2012). Individual differences in late bilinguals' L2 phonological processes: from acoustic-phonetic analysis to lexical access. *Learn. Individ. Differ.* 22, 680–689. doi: 10.1016/j.lindif.2012.05.005
- Draxler, C., and Jänsch, K. (2004). *SpeechRecorder. A Universal Platform Independent Multichannel Audio Recording Software*. Version 3.8.0. München.
- Eger, N. A., and Reinisch, E. (2019). The role of acoustic cues and listener proficiency in the perception of accent in non-native sounds. *Stud. Second Lang. Acquisit.* 41, 179–200. doi: 10.1017/S0272263117000377
- Faris, M. M., Best, C. T., and Tyler, M. D. (2016). An examination of the different ways that non-native phones may be perceptually assimilated as uncategorized. *J. Acoust. Soc. Am.* 139, EL1–EL5. doi: 10.1121/1.4939608
- Flege, J. E. (1995). "Second language speech learning: theory, findings, and problems," in *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, ed. W. Strange (Timonium, MD: York Press), 233–276.
- Gendrot, C., Kühnert, B., and Demolin, D. (2015). "Aerodynamic, articulatory and acoustic realization of French /ʁ/," in *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS'15)*, Glasgow, United Kingdom.
- Gibson, E. J. (1991). *An Odyssey in Learning and Perception*. Cambridge, MA: Bradford Books (MIT Press).
- Gibson, J. J. (1966). *The Senses Considered as Perceptual Systems*. Boston, MA: Houghton Mifflin.
- Guion, S. G., Flege, J. E., Akahane-Yamada, R., and Pruitt, J. C. (2000). An investigation of current models of second language speech perception: the case of Japanese adults' perception of english consonants. *J. Acoust. Soc. Am.* 107, 2711–2724. doi: 10.1121/1.428657
- Harnsberger, J. D. (2001). On the relationship between identification and discrimination of non-native nasal consonants. *J. Acoust. Soc. Am.* 110, 489–503. doi: 10.1121/1.1371758
- Hausser, J., and Strimmer, K. (2009). Entropy inference and the James-Stein estimator, with application to nonlinear gene association networks. *J. Mach. Learn. Res.* 10, 1469–1484. doi: 10.1145/1577069.1755833
- Hothorn, T., Bretz, F., and Westfall, P. (2016). Simultaneous Inference in General Parametric Models vignette ("generalsiminf", package = "multcomp").
- Hualde, J. I. (1991). *Basque Phonology*. London, UK: Routledge.
- Hualde, J. I. (2005). *The Sounds of Spanish with Audio CD*. Cambridge, UK: Cambridge University Press.
- Hualde, J. I., Ollarrea, A., Escobar, A. M., and Travis, C. E. (2010). *Introducción a la Lingüística Hispánica, 2nd edn.* Cambridge, UK: Cambridge University Press.
- Kohler, K. (1995). *Einführung in die Phonetik des Deutschen (Grundlagen der Germanistik, Band 20)*. Berlin, Germany: Erich Schmidt Verlag.
- Lancaster, A., and Gor, K. (2016). Abstraction of phonological representations in adult nonnative speakers. *Proc. Linguist. Soc. Am.* 1, 24–21. doi: 10.3765/plsa.v1i0.3725
- Llompert, M., and Reinisch, E. (2019a). Imitation in a second language relies on phonological categories but does not reflect the productive usage of difficult sound contrasts. *Lang. Speech* 62, 594–622. doi: 10.1177/0023830918803978
- Llompert, M., and Reinisch, E. (2019b). The robustness of lexical representations in a second language relates to phonetic flexibility for difficult sound contrasts. *Biling. Lang. Cogn.* 22, 1085–1100. doi: 10.1017/S1366728918000925
- Llompert, M., and Simonet, M. (2018). Unstressed vowel reduction across majorcan catalan dialects: production and spoken word recognition. *Lang. Speech* 61, 430–465. doi: 10.1177/0023830917736019
- Martin, A., and Peperkamp, S. (2017). Assessing the distinctiveness of phonological features in word recognition: prelexical and lexical influences. *J. Phonet.* 62, 1–11. doi: 10.1016/j.wocn.2017.01.007
- Mitterer, H., and Reinisch, E. (2017). Surface forms trump underlying representations in functional generalisations in speech perception: the case of German devoiced stops. *Lang. Cogn. Neurosci.* 32, 1133–1147. doi: 10.1080/23273798.2017.1286361
- Mitterer, H., Reinisch, E., and McQueen, J. M. (2018). Allophones, not phonemes in spoken-word recognition. *J. Mem. Lang.* 98, 77–92. doi: 10.1016/j.jml.2017.09.005
- Mitterer, H., Scharenborg, O., and McQueen, J. M. (2013). Phonological abstraction without phonemes in speech perception. *Cognition* 129, 356–361. doi: 10.1016/j.cognition.2013.07.011
- Morey, R. D. (2008). Confidence intervals from normalized data: a correction to cosineau (2005). *Tutor. Quant. Methods Psychol.* 4, 61–64. doi: 10.20982/tqmp.04.2.p061
- Navarro-Tomás, T. (1916). *Las vibraciones de la "rr" española*. Española: Revista Filol 3, 166–168.
- Peirce, J. W. (2007). PsychoPy—psychophysics software in python. *J. Neurosci. Methods* 162, 8–13. doi: 10.1016/j.jneumeth.2006.11.017
- Pinheiro, J. C., and Bates, D. M. (2000). "Linear mixed-effects models: basic concepts and examples," in *Mixed-Effects Models in S and S-Plus* (New York, NY: Springer), 3–56. doi: 10.1007/978-1-4419-0318-1_1
- Pitt, M. A., Dille, L., and Tat, M. (2011). Exploring the role of exposure frequency in recognizing pronunciation variants. *J. Phonet.* 39, 304–311. doi: 10.1016/j.wocn.2010.07.004
- R Core Team (2017). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*. Vienna, Austria. Available online at: <https://www.R-project.org/>.
- Ranbom, L. J., and Connine, C. M. (2007). Lexical representation of phonological variation in spoken word recognition. *J. Mem. Lang.* 57, 273–298. doi: 10.1016/j.jml.2007.04.001
- Schiller, N. O. (1999). "The phonetic variation of German /r/. variation und stabilität in der wortstruktur," in *Untersuchungen zu Entwicklung, Erwerb und Varietäten des Deutschen und anderen Sprachen* (Hildesheim, Zürich, New York, NY: Georg Olms Verlag), 261–287.
- Sebastián-Gallés, N., Echeverría, S., and Bosch, L. (2005). The influence of initial exposure on lexical representation: comparing early and simultaneous bilinguals. *J. Mem. Lang.* 52, 240–255. doi: 10.1016/j.jml.2004.11.001
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 379–423. doi: 10.1002/j.1538-7305.1948.tb01338.x
- Sumner, M. (2015). The social weight of spoken words. *Trends Cogn. Sci.* 19, 238–239. doi: 10.1016/j.tics.2015.03.007
- Sumner, M., Kim, S. K., King, E., and McGowan, K. (2014). The socially-weighted encoding of spoken words: a dual-route approach to speech perception. *Front. Psychol.* 4:1015. doi: 10.3389/fpsyg.2013.01015
- Sumner, M., and Samuel, A. G. (2005). Perception and representation of regular variation: the case of final /t/. *J. Mem. Lang.* 52, 322–338. doi: 10.1016/j.jml.2004.11.004

- Trubetzkoy, N. S. (1977). *Grundzüge der Phonologie, 6th Edn.* Göttingen: Van den Hoeck & Ruprecht.
- Tyler, M. D., Best, C. T., Faber, A., and Levitt, A. G. (2014). Perceptual assimilation and discrimination of non-native vowel contrasts. *Phonetica* 71, 4–21. doi: 10.1159/000356237
- Ulbrich, C., and Ulbrich, H. (2007). “The realisation of /r/ in Swiss German and Austrian German,” in *Proceedings of the 16th International Congress of Phonetic Sciences (Saarbrücken) 1761–1764.*
- Ulbrich, H. (1973). Zur kodifizierung der R-Aussprache im Siebs. *STUF Lang. Typol. Univ.* 26, 120–133. doi: 10.1524/stuf.1973.26.16.120
- Van Bezooijen, R. (2005). Approximant /r/ in Dutch: routes and feelings. *Speech Commun.* 47, 15–31. doi: 10.1016/j.specom.2005.04.010
- van Leussen, J.-W., and Escudero, P. (2015). Learning to perceive and recognize a second language: the L2LP model. *Front. Psychol. Lang. Sci.* 6:1000. doi: 10.3389/fpsyg.2015.01000
- Wiese, R. (1996). *The Phonology of German. The Phonology of the World's Languages.* Oxford: Clarendon Press.
- Wiese, R. (2003). The unity and variation of German /r/. *Zeitschrift für Dialektol. Linguist.* 70, 25–43. doi: 10.2307/40504887

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Reinisch, Juhl and Llompart. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.