Shea, C. & Lira, C. (2020). L2 production of Brazilian nasal vowels in online dyads: Does a native speaker model make a difference? In O. Kang, S. Staples, K. Yaw, & K. Hirschi (Eds.), *Proceedings of the 11th Pronunciation in Second Language Learning and Teaching conference*, ISSN 2380-9566, Northern Arizona University, September 2019 (pp. 186–196). Ames, IA: Iowa State University.

L2 PRODUCTION OF BRAZILIAN NASAL VOWELS IN ONLINE DYADS: DOES A NATIVE SPEAKER MODEL MAKE A DIFFERENCE?

Christine Shea, Department of Spanish and Portuguese, University of Iowa
Cristiane Lira, Department of Spanish and Portuguese, University of Iowa

We examine the production of nasal/nasalized vowels in Portuguese by native speakers of English. Participants interacted in an online spoken conversation with Brazilian Portuguese speakers. Words with the target segments were extracted from the dialogues and categorized in terms of their distance from a native speaker production of the same target type: immediately following a native-speaker example, two turns after a native speaker example, more than three turns after a native speaker example. Twenty-seven Brazilian Mechanical Turk workers judged the productions and rated them on a scale of 1(heavily-accented) to 6 (native-like). Results show that the type of nasal had a greater effect on accentedness judgments than temporal proximity to a native speaker production.

INTRODUCTION

This study investigates variability in L1 English/second language (L2) Brazilian Portuguese (BP) production of BP nasal/nasalized vowels. We focus on how L1 speakers adjust (or not) their production of lexical items with the target vowels in relation to a previous production of the target vowel by a native BP speaking interlocutor. This process is known as *convergence* (also known as accommodation, imitation, alignment, see work by Pardo, 2006, *inter alia*) and captures the ways in which speakers adapt their linguistic behaviour according to characteristics of their interlocutor (such characteristics can be speech-related or extra-linguistic, such as social proximity, gender, or impressions of shared solidarity).

To account for the multiple phenomena related to convergence, Pickering and Garrod (2013) propose that perception and production are closely coupled. Specifically, they suggest that during a communicative interaction, speakers construct forward models of their gestures before they actually produce speech, and perceivers covertly imitate those actions and then create forward models of those actions. In other words, perception and production are joint actions that take advantage of prediction and constantly updated interactions. Under this approach, language perception and production rely upon the same underlying representations, and talkers and listeners can become aligned over the course of a conversation, in a process known as 'input-output' coordination. Phonetic representations activated as part of comprehension are then available to automatically shape subsequent language production

The relevance of this model to the present study lies in the way it can account for the tight coupling between perception and production and also for potential temporal proximity effects – that is, the more recently a speaker hears an exemplar of a particular word or sound, the stronger the representation and greater likelihood that the speaker will imitate what they just heard, through a process of covert imitation. Covert imitation occurs when the speaker is not aware of how their speech is approximating that of their interlocutor and can be contrasted with more conscious imitation, i.e., when speakers are consciously trying to imitate a previous model of speech.

One issue that is not explicitly covered in Garrod and Pickering's model, however, is how it might apply to bilingual – and particularly unbalanced bilingual – contexts. Specifically, there is a great deal of evidence across all levels of language knowledge that later-in-life bilinguals need to suppress their dominant language when producing and processing the other. This raises questions for models such as Garrod and Pickering's that assume covert imitation and prediction as underlying the perception-production process. In the case of L2 production, the ease of prediction relies on relatedness and also familiarity (Festman, 2013). Familiarity can take the shape of articulatory routines or perceptual classification. Predictions are easier and more likely to be correct when interlocutors are equally proficient in the language being used. Thus, it is possible that while phonetic convergence may occur in the case of L2 speakers, it will be modulated by the specific segments and words being produced and perceptual factors (i.e., if the L2 interlocutor cannot perceive the sound, she will not be able to imitate it). Given this, we predict an interaction between the segment produced and its timing relative to the previous production by a native BP speaker in the same interaction.

To test our hypotheses, we use a corpus of online, naturalistic conversational data to investigate the pronunciation of Brazilian Portuguese nasal vowels, nasalized vowels and nasal diphthongs by L1 English/L2 Portuguese students. We extracted a sample of the L2 productions and presented them to native Brazilian Portuguese listeners via Mechanical Turk, who rated them on a scale of 1 (heavy English accent) to 6 (very little English accent). Productions were coded for vowel type (nasalized, nasal, nasal diphthong) and repetition (immediately following a native speaker production, two or three turns after a native speaker production or unprompted).

Brazilian Portuguese Nasal and Nasalized Monophthongs and Nasal Diphthongs

The BP inventory of vowel phonemes is relatively large, consisting of seven oral monophthongs /i e ɛ a ɔ o u/ and five nasal monophthongs [ĩ ẽ ɐ̃ õ ũ]. In addition to the monophthongs, syllable rimes may have as many as 13 oral diphthongs, four oral triphthongs, five nasal diphthongs, and two nasal triphthongs (Cristófaro-Silva, 2007; Whitlam, 2017).[2]

Nasality is one of the most characteristic and controversial aspects of Portuguese. While it is commonly accepted that Portuguese has contrastive nasal vowels, there is a lack of absolute consensus across the phonological literature regarding their status. Phonologically, studies have proposed that nasality in a vowel is a binary feature that can emerge as an output of a nasalization rule, in which the nasalized vowel shares the property [+nasal] with the following nasal consonant. This is the case of a nasalized vowel, as in [ˈkɐ̃.mɐ] *cama* 'bed'. It is regarded as a coarticulation process, and it is rather different from nasality from phonemic nasal vowels, in which the vowel is also nasalized as part of a nasalization process, but the conditioning nasal consonant may be deleted on the derivation process, yielding a full nasal vowel, as in [ˈkɐ̃^{n}.tʊ] *canto* 'sing' (Seara, 2000).

[2] Given space restrictions, we cannot do justice to the vast literature on nasalization in Brazilian Portuguese. Readers are encouraged to consult work by Marques (2018), Seara (2000), among others for details and discussion of the issues.

In BP, there are four nasal diphthongs: /ẽɪ̯, õɪ̯ ɐ̃ʊ̯, ɐ̃ɪ̯/ that can occur in in word-final position or in word-final with a nasal orthographic symbol following it (see Table 1). Across dialects, there is a great deal of variability in their realization.

Table 1 presents examples of the nasal and nasalized monophthongs and nasal diphthongs found in Brazilian Portuguese:

Table 1

Examples of Brazilian nasal/nasalized monophthongs and diphthongs (Porter, 2015)

Nasalized monophthongs [V.N] *(noncontrastive)*			**Nasal monophthong [$V^{N}C$]** *(contrastive)*			**Nasal diphthong [ṼG]** *(contrastive)*		
ac<u>i</u>ma	[asĩma]	‘above’	*banco*	[ˈbɐ̃^{n}kʊ]	‘bank’	*pão*	[pɐ̃ʊ̯]	‘bread’
m<u>a</u>ta	[mãta]	‘kill’	*tempo*	[ˈtẽ^{m}pʊ]	‘weather’	*mãe*	[mɐ̃ɪ̯]	‘mom’
c<u>a</u>ma	[kɐ̃ma]	‘bed’	*pinta*	[ˈpĩ^{n}ta]	‘paint’	*também*	[tẽ^{m}bɐ̃ɪ̯]	‘also’
moto	[mõto]	‘motorcycle’	*sombra*	[ˈsõ^{m}brɐ]	‘shadow’	*limões*	[limõɪ̯s]	‘lemons’
pena	[pẽna]	‘pity’	*mundo*	[ˈmũ^{n}dʊ]	‘world’			
como	[kɔ̃mo]	‘eat’						
f<u>u</u>ma	[fũma]	‘smoke’						

Phonetically, nasal vowels can be followed by a nasal resonance, called nasal appendix (Moraes, 2013), which, depending upon the analysis, may be considered the remnant of a nasal consonant that nasalizes the vowel (Marques & Scarborough, 2017). The degree of nasality between phonemically nasal ([ˈkɐ̃^{n}.tʊ]) and coarticulatory nasalized vowels ([ˈkɐ̃.mɐ]) can vary, depending on the degree of the velopharyngeal port opening. Aerodynamic studies show that nasal vowels have a higher nasal airflow than nasalized vowel (Medeiros, 2011).

L2 Production of Brazilian Portuguese Nasal and Nasalized Vowels

We will briefly present some of the learning challenges faced by L1 English/L2 Portuguese learners when acquiring the nasal vowels in their L2. English does not have nasal vowels and while nasalized vowels occur (e.g., *man* [mæ̃n]), the degree of nasality is less than that of Portuguese and nasality is never used contrastively. For English speakers, nasal vowels ($V^{N}C$, e.g., *canto* ‘sing’) in Portuguese are often produced with pronounced nasal closure before the consonant and a smaller degree of nasalization on the vowel results, a direct transfer from L1 English words such as *tandem.* In English, a coda-final nasal consonant is fully pronounced and can close syllables. When L1 English speakers are learning Portuguese, they tend to transfer the full nasal closure to Portuguese. That is, when there is a nasal consonant in the input (reinforced by spelling), native English speakers pronounce it, as per the phonology of their native language. It is necessary for L1 English/L2 Portuguese speakers to learn that in fact, there is no full nasal closure in their L1. Instead, there is a small (and at times non-existent) nasal closure at the end of the vowel, before the following non-nasal consonant (see discussion above). In the case of nasalized vowels, the situation is distinct. Since nasalization is allophonic in English, speakers of this language may not be aware that they are producing it, which may mean they are unaware of the similar nasalization process that occurs in Portuguese. Finally, the nasal diphthong does exist in English (e.g., ‘mine’, or ‘noun’) but it is allophonic and only occurs in syllables closed by nasal consonants.

In Portuguese, the nasal diphthong can occur in open syllables (*mãe* 'mom') and, importantly as the result of morphological inflection: the plural form for various nouns in Portuguese requires nasal diphthongs (e.g., *limões*, 'lemons'). Given this brief comparison between English and Portuguese vowel nasalization, we propose that the order of difficulty for producing Portuguese nasal monophthongs and diphthongs and nasalized monophthongs, will be the following: nasalized monophthongs, nasal monophthongs > nasal diphthongs.

Specifically, the nasal diphthongs will represent the greatest challenge for L1 English learners while the nasalized and nasal monophthongs will be produced with greater accuracy (as per native speaker judgments).

Hypotheses

We focus on two independent variables in this study: vowel type and L2 production across time in relation to the native speaker production. Our predictions are the following:

1. Repetition: immediate repetitions will be produced with greater native-like precision, followed by the two-turn repetition and finally, the unprompted productions overall, for all vowels.
2. Nasal type: the nasal diphthongs will be produced with less native-like precision than the nasal vowels and the nasalized vowels, across all repetitions.
3. Building on hypotheses 1 and 2, we predict that the unprompted nasal diphthong will receive the lowest ratings (i.e., strongest English accent) by the native BP listeners.

METHODS

Data

The tokens provided to the BP listeners were extracted from the Multimodal Teletandem Corpus (MulTeC, Aranha, Luvizari-Murad & Moreno, 2015), compiled by researchers at UNESP (São Paulo State University, São José do Rio Preto). Teletandem is a virtual conversational exchange in which students help each other learn their respective native languages using webcam transmission. The speakers were L1 English/L2 Portuguese undergraduate students studying at American institutions and L1 BP/L2 English learners studying at Brazilian universities. The speakers we used were at an advanced beginner/low-intermediate level of Portuguese.

Procedure

The conversational dyads were selected from those available through MulTeC. The corpus managers granted access to dyads that included speakers from the proficiency levels we sought (based upon their semester enrollment in L2 Portuguese), starting at from the earliest stages of Teletandem operation (2015). Subsequently, the first author listened to the dyads and determined which ones were of sufficient acoustic quality to serve as data for the study. Given that the goal of Teletandem is not to provide phonetic data, the quality of the recordings at times was severely compromised. Furthermore, the participants are instructed to have naturalistic conversations, which meant that it was not possible to extract the same words across different speakers or across different repetition conditions. A total of 42 speaker dyads were judged for recording quality, from which 26 were determined appropriate the study. From these, we extracted the highest quality items, and where there were two items that were of equal quality, we used the word that was most similar to the native speaker production (in terms of the nasal/nasalized vowel). Where that was not possible, we selected the word that referred to a noun (rather than a verb or adverb).

The recordings from the videos were extracted and converted to .wav files. The second author (native BP speaker from Sãu Paolo) then listened to the sound files and marked BP speaker nasal productions and L1 English/L2 Portuguese nasal productions. The first author then coded the L2 productions for vowel type and interaction.

The tokens were presented via Amazon Mechanical Turk to 27 native BP listeners, living in Brazil.[3] Raters were told they were going to hear L1 English speakers produce words in Portuguese, which they had to rate on a six-interval scale, provided to them on the screen (1=very strong English accent, 6=very little English accent). There were 18 tokens in total, two for each orthogonal combination of vowel and repetition, produced by twelve different speakers.

RESULTS

Figure 1 shows the count of each rating value for the vowel-by-repetition combinations. The rows refer to the repetition condition and the vertical columns refer to the different target vowels. As can be seen, the nasal diphthongs received a greater number of lower ratings overall than did the two other target vowels. The noncontrastive nasalized vowels (*mente*) in the unprompted condition received the greatest number of ratings above three.

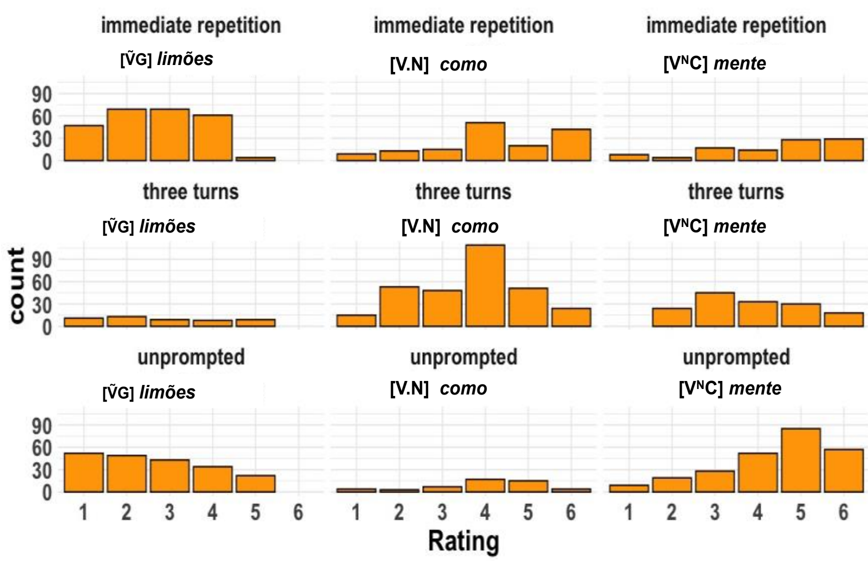

Figure 1. Count of ratings for vowel-by-repetition tokens

As the ratings were coded into ordinal variables (on a 6-point scale), we used mixed ordinal logistic regression to examine the effects of vowel type and repetition on ratings by the BP listeners. In ordinal logistic regression the dependent variable is usually some sort of rating scale, which means

[3] AMT allows users to specify the IP address of the workers. We limited ours to Brazilian IPs.

there is an explicit ordering in the categories (from highest rating to lowest). Ordinal logistic regression allows us to calculate the probability that a particular variable will receive a specific score, or a score below a certain threshold by using cumulative events for the log of the odds computation. Unlike simple logistic regression, ordinal logistic models consider the probability of an event and all the events that are below it in the ordered hierarchy. In the present study, the probability of each score on the 1-6 scale are the events under analysis. For example, by using ordinal logistic regression, we can calculate the probability that a nasal diphthong produced immediately after a native speaker production will receive a rating of '4' or, alternatively, the probability that this category of tokens will receive a rating higher than '3' (i.e., of '4', '5' or '6').

The maximum fitted model included random slopes for judge and word and fixed effects for nasal type and repetition, entered as an interaction. The logit link function was used, which is equivalent to the proportional odds model. The models were fitted with the *clmm2* function from the ordinal package (Christensen, 2011) in R. We calculated the interrater reliability for ordinal ratings using Krippendorff's alpha which indexes the overall agreement among the raters for the samples. The results revealed a value of .72, which corresponds to high reliability.

The odds of higher ratings across the three times reached significance only for the unprompted vs. immediate productions, where the odds of receiving a higher score for the immediate repetition were 4.2 times greater than for the immediate repetition (p=.03). The other coefficients did not reach significance. For nasal type, the odds of receiving a higher score were 10.8 times greater for the V.N items than for the nasal diphthong (p<.001) and the odds of receiving higher scores for the $V^{N}C$ items were 13.01 times greater than for the nasal diphthongs (p<.001). The interaction between nasal type and repetition was not significant (p=.101).

Using the function package *ggpredict* in R, it is possible to get the overall probabilities of each rating level for the model and also the probabilities for each rating across each level of the predictor variables. To determine the overall odds of receiving a particular rating score, we can look at the threshold coefficients, which, when converted to odds ratios, give us the 'cut points' between the levels of the dependent variable, in this case, the rating scale. The threshold coefficient for a value of 3 or less overall (across the entire model) is 2.5, which can be interpreted as the likelihood of receiving a score of 1, 2, or 3 rather than a 4, 5, or 6. In comparison, the probability of receiving a rating of 1, 2, 3, or 4 rather than a 5 or 6 is 9.61, or almost four times greater, indicating that, overall, the nonnative speakers were rated four times more often with scores of 4 or less ('a slight English accent').

In addition to the overall probabilities of each rating for the model, it is possible to obtain the probabilities for each rating type across the levels of each predictor variable using the *plogis* function in R. Figures 2 and 3 present these probabilities for repetition and nasal type.

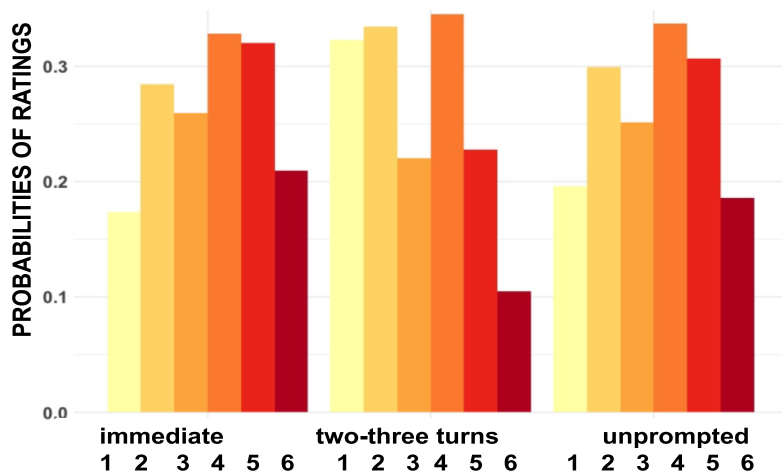

Figure 2. Probabilities of ratings for repetition

As can be seen in Figure 2, the probabilities of receiving specific ratings across the three repetition conditions do not pattern in the predicted fashion (Hypothesis 1). Specifically, the unprompted productions, overall, do not show a higher probability of receiving a lower rating than the immediate repetitions. For example, the probability of receiving a rating of 6 (very little English accent) ranges from .21 for the immediate repetition condition to .12 for two-three turns and .18 for the unprompted condition. And the probability of receiving a rating of three or lower for the immediate repetition condition is .67 while for the two-three turns it is .74 and for the unprompted productions it is .71. For the same levels of repetition, the probabilities of receiving a rating of 5 or higher are .20, .31and .29, respectively.

Figure 3 shows the probabilities of receiving specific ratings across the nasal type condition. The prediction that the nasal diphthong would receive the lowest rating is borne out by the data. The probability of the nasal diphthong receiving a rating of three or lower is .71 while for the V.N tokens, the probability is .18 and for the $V^{N}C$ tokens, it is .23. For ratings of five or six, the probabilities are .07, .21 and .28, respectively. These results support Hypothesis 2, which predicted that the nasal diphthongs would receive the lowest ratings overall.

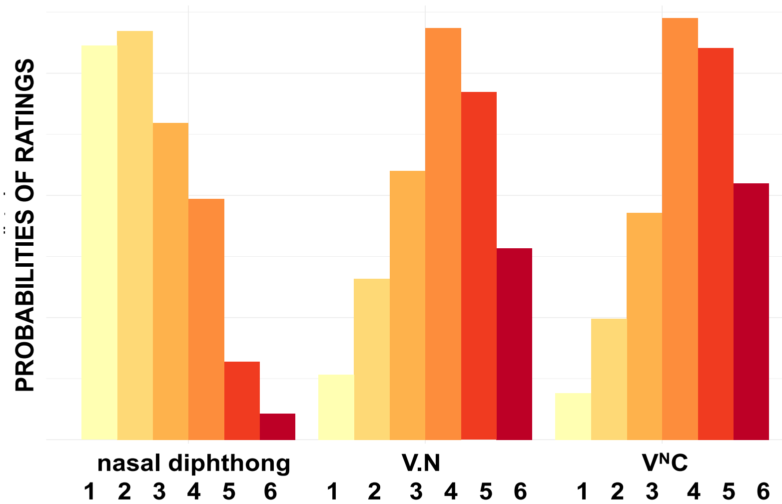

Figure 3. Probabilities of ratings for nasal type

In Figure 4 we present the probabilities by vowel type and repetition. The final hypothesis predicted an interaction between repetition and nasal type. There were no significant interactions observed in the data, leading us to reject Hypothesis 3.

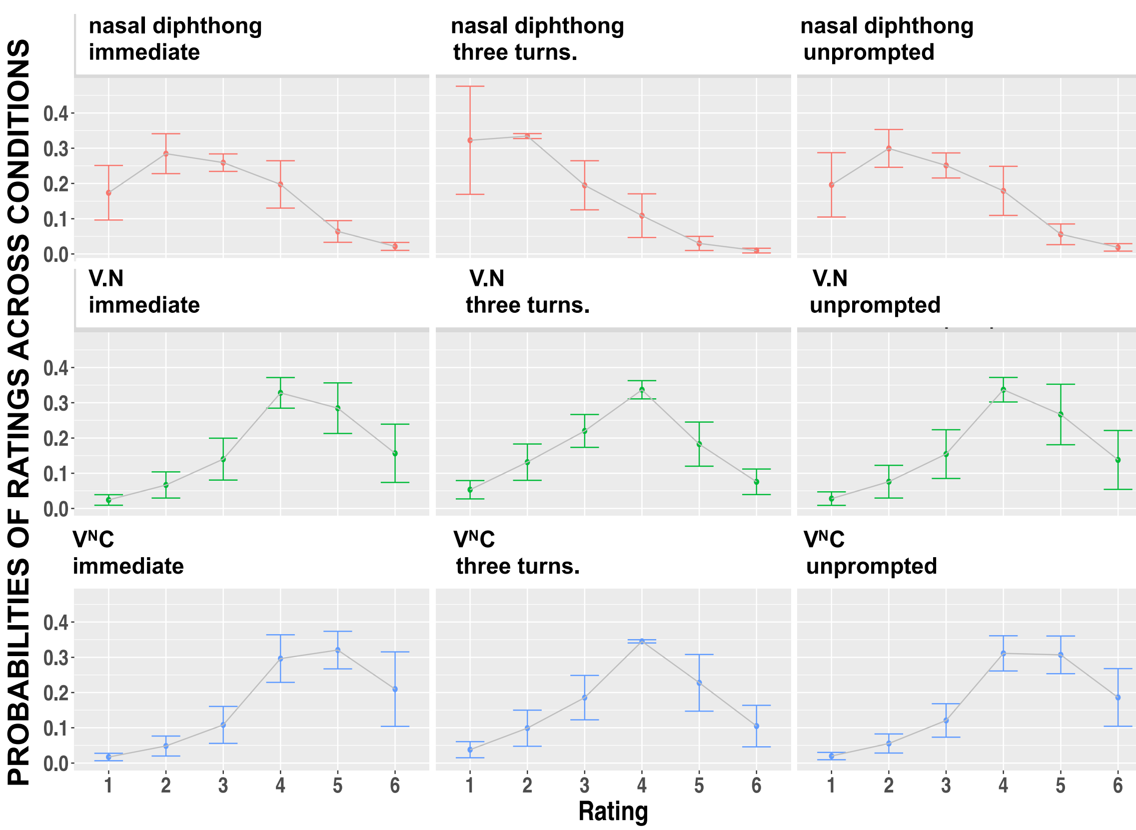

Figure 4. Probabilities of ratings across nasal type and repetition.

DISCUSSION

In this study we investigated whether hearing a native-speaker model of a difficult L2 segment would lead to more native-like productions of that segment. We analyzed naturalistic productions of nasalized and nasal monophthongs and diphthongs in Brazilian Portuguese by L1 English/L2 Brazilian Portuguese learners to test whether the timing of their productions with respect to a native-speaker model modulated the perceived accentedness of their tokens. We hypothesized that productions occurring immediately after a native speaker token would be more nativelike while those produced unprompted would be least native-like. We also predicted that the nasal diphthong would be judged as least native-like, followed by the nasal monophthongs and finally, the nasalized monophthongs. Judges were native Brazilian Portuguese listeners recruited through Amazon Mechanical Turk, who were asked to rate the productions on a scale of 1 (heavy English accent) to 6 (very little English accent).

The results did not support the repetition hypothesis. The unprompted productions were not significantly less likely to receive evaluations of 1 or 2 than the two turn repetitions or the immediate repetitions. The nasal type hypothesis, on the other hand, was robustly supported. The nasal diphthongs were consistently more likely to receive ratings of 3 or less than the other nasal types. This results most likely washed out any possible effects for the timing variable.

What do these results mean for models such as Pickering and Garrod (2013) which propose a close coupling between perception and production? We argue that second language speech requires its own consideration, given the inherent difficulties with learning new perceptual and articulatory routines. Specifically, coordination among speakers may be modulated by perceived proximity among interlocutors and also, of course, by the segments being produced. In the case of the nasal vowels of BP, English L1 speakers typically require a great deal of practice and input before they can reach native-like production accuracy. We may have observed different results if we had chosen a less difficult target (e.g., vowel quality in non-nasal vowels).

This leads us to one of the drawbacks of the present study. Because the corpus from which we drew our data (MulTeC, Aranha et al., 2015) was not designed for phonetics/phonology-based research, it was not possible to control for the type of dialogue produced and the quality of the recordings did not permit a more detailed acoustic analysis of the tokens themselves. As well, since the lexical items were taken from authentic, uncontrolled interactions, it was not possible to use the same words across all speakers. Some of the speakers may have attempted words that were less frequent than others (less perceptual/production experience), which may also have affected the results reported here. Finally, we were forced to rely upon the judgments of our Brazilian AMT workers and their interpretation of the instructions provided to them on the screen.

In future work, we would like to examine the production of BP nasal vowels by native English speakers under more ideal recording conditions and perhaps with more advanced speakers, who have come closer to mastery of the target segments.

ACKNOWLEDGMENTS

Thank you to the audience at PSLLT 2019 for thoughtful comments on our study. We would also like to thank Dr. Solange Aranha (UNESP) for granting us access to MuTeC. Finally, we would like to extend our thanks as well to all the students in Brazil and the United States who gave permission for their data to be included in the corpus.

ABOUT THE AUTHORS

Dr. Christine Shea is an Associate Professor of Spanish Linguistics in the Spanish and Portuguese Department at the University of Iowa. Her research focuses on the acquisition of second language phonetics, phonology and speech perception. Her research has appeared in *Second Language Research*, *Studies in Second Language Acquisition* and the *Modern Language Journal*.
christine-shea@uiowa.edu

Dr. Cristiane Lira is the director of the Portuguese program at the University of Iowa. Her research focuses on the representation of female guerrilla warfare fighters in Brazil and Argentina during the last dictatorship in both countries (1964-1985). Her research has appeared in *Chasqui* and she has a forthcoming article in *Revista Iberoamericana* (2020). She recently published *Ponte para o poente (Bridge to the sunset)* and *No país da infância* (*My childhood country)* as part of a collection of books organized by the women's collective Mulherio das Letras (Women of Letters).
cris-lira@uiowa.edu

REFERENCES

Aranha, S., Luvizari-Murad, L.H., Moreno, A.C. (2015) A criação de um banco de dados para pesquisas sobre aprendizagem via teletandem institucional integrado (Ttdii). *Revista (Con)textos Linguísticos 9*(12), 274–293.

Christensen, R.H.B. (2011) A tutorial on fitting cumulative link mixed models with clmm2 from the ordinal package. http://cran.r-project.org/web/packages/ordinal/vignettes/clmm2_tutorial.pdf

Cristófaro Silva, T. (2007). Fonética e fonologia do português (9th ed). São Paulo: Editora Contexto.

Festman, J. (2013). The complexity-cost factor in bilingualism. *Behavioral and Brain Sciences*, *36*(4), 355-356.

Marques, L. F. (2018). *Perception of Vowel Nasality Contrast in Brazilian Portuguese*. PhD Dissertation, University of Colorado.

Marques, L., & Scarborough, R. (2017). Perception and Acoustics of Vowel Nasality in Brazilian Portuguese. *INTERSPEECH*, 616-620.

Medeiros, B. (2011). Nasal coda and vowel nasality in Brazilian Portuguese. In M. Scott (Ed.) *Selected Proceedings of the 5th Conference on Laboratory Approaches to Romance Phonology,* Boston:Cascadilla Proceedings Project, pp. 33-45.

Moraes, J. (2003). A nasalidade vocálica no português do Brasil e no português do Portugal. *Actas del XXIII Congreso Internacional de Lingüística y Filología Románica*. Tübingen: Max Niemeyer Verlag, pp 235-44.

Moraes, J. (2013). Produção e percepção de vogais nasais in M. Abaurre (Ed.) *Gramática Do Português Culto Falado No Brasil: A ConstruçãoFonológica Da Palavra*. São Paulo, Brazil: Contexto, pp. 95-112.

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America*, *119*(4), 2382-2393.

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and brain sciences*, *36*(4), 329-347.

Porter, D. (2015). Progressive Vowel Nasalization in Brazilian Portuguese: A Preliminary Analysis. *Spanish and Portuguese Review*, 1-19.

Seara, I. (2000). *Analysis acoustic-perceptual of nasality of the vowels in the Brazilian Portuguese.* Doctoral dissertation, Federal University of Santa Catarina, Brazil.

Whitlam, J. (2017). *Modern Brazilian Portuguese grammar: a practical guide*. Routledge.