

THE INFLUENCE OF METALINGUISTIC KNOWLEDGE OF SEGMENTAL PHONOLOGY ON THE PRODUCTION OF ENGLISH VOWELS BY BRAZILIAN UNDERGRADUATE STUDENTS

Ronaldo Manguiera Lima Júnior*
Universidade Federal do Ceará, BRA

Abstract

This article presents data on the production of English vowels [i ɪ ε æ u ʊ] by Brazilian English Language Teaching (ELT) undergraduate students before and after taking a course on English Segmental Phonology. Brazilian learners tend to assimilate the contrasts present in [i ɪ], [ε æ] and [u ʊ] into the prototypical categories of Brazilian Portuguese [i], [ε] and [u], respectively. Thus, this article investigates the influence of receiving explicit metalinguistic instruction of English segmental phonology on the production of the target pairs of vowels. The data analysis is of acoustic nature (spectral quality), and the results show that some learners created new phonetic categories for the English vowels after receiving the metalinguistic instruction.

Keywords: phonological acquisition; acoustic analysis; English as a Foreign Language; explicit instruction

1 Introduction

Having a minimum control of the pronunciation of a second/foreign language (L2) is required for oral communication. Accurate production of both segments and prosodic elements are necessary for intelligibility (e.g. Celce-Murcia, Brinton, & Goodwin, 2010; Morley, 1994; Pennington, 1996, 1998), which, roughly explained, refers to how much your interlocutor understands of your speech; and for comprehensibility (Parrino, 1998; Singleton & Ryan, 2004), which is related to the effort your interlocutor needs to employ in order to understand your speech. Therefore, it goes without saying that pronunciation instruction, even though very often neglected by language teachers for various reasons, should be a constant part of L2 pedagogy (Silveira, 2004; Tomlinson, 2005; Yule & Macdonald, 1994).

As a result, non-native English speaking undergraduate students of English Language Teaching (ELT) should be especially interested in improving

their pronunciation of the L2 for two basic reasons: a) they will be a model for their students, and probably the greatest source of input and exposure to the language for most learners; and b) with enough knowledge of the L2 sound system, these future English teachers will be more confident and empowered to engage in pronunciation instruction. That is the reason most ELT undergraduate programs in Brazil include at least one mandatory course on English phonetics and/or phonology. This article, in turn, aims to investigate how Brazilian ELT undergraduate students may improve their production of English vowels [i ɪ ε æ u ʊ], which are particularly challenging for Brazilian Portuguese speakers, after taking a graduate-level course on English Segmental Phonology.

The six vowels in focus are particularly challenging for Brazilian learners due to the natural difficulty to perceive and produce sounds of an L2 which are not present, or not contrasted, in the learner's native language (L1). This difficulty does not exist because people lose their ability to learn new sounds at a certain

* Holds a doctorate in Linguistics and master's in Applied Linguistics, both from the University of Brasília (UnB). He is currently a professor at the Department of English Studies at the Federal University of Ceara, Brazil. His main research interest is in L2 phonological acquisition. His e-mail address is ronaldojr@letras.ufc.br



age, but because people learn the phonological system of their L1 so well. When acquiring their L1, one needs to learn how to accommodate the variation inherent to the acoustic signal into prototypical phonological categories of their L1 so that communication can take place, and the brain does so by taking statistics of the input and assigning exemplars to the corresponding categories (Bybee, 2003; Cristófaró Silva, 2003; P. Kuhl, 1991; P. K. Kuhl, 1993; P. K. Kuhl et al., 2008; Leather, 2003; Pierrehumbert, 2001, 2003).

Therefore, identifying and, consequently, producing L2 sounds that are acoustically very close to a sound of the L1 become more challenging. The L2 sounds that are phonetically closer to sounds of the L1 are the hardest for L2 learners to perceive and produce, since they tend not to (initially) perceive them as different, assimilating them to the prototypical phonological categories of their L1 (Flege, 1995, 1999, 2007). This is the case with English vowels [i ɪ], [ɛ æ] and [u ʊ], which tend to be assimilated by Brazilian learners into the prototypical categories of Brazilian Portuguese [i], [ɛ] and [u], respectively (Bion, Escudero, Rauber, & Baptista, 2006; Lima Jr, 2015; Nobre-Oliveira, 2007; Rauber, 2006). That is the reason those six vowels are the ones in focus in this article.

Assuming that the process of L2 acquisition is a complex dynamic system (e.g. De Bot, 2008; De Bot, Lowie, & Verspoor, 2007; Larsen-Freeman, 1997; Lima Jr, 2013), the prototypical categories created for communication in the L1 act as attractor states for the L2. Attractors are states of temporary accommodation of a complex dynamic system; it is where the system finds temporary stability amidst chaos. The very fact that attractor states are temporary reinforces the dynamic nature of such system, meaning that the system is constantly moving from one attractor state to another. Some attractor states require more energy for the system to move away from, but they are all potentially temporary in nature.

This means that Language Acquisition would be more accurately described as Language Development,¹ due to its dynamic, never-ending change in time as the system moves through different attractor states. Even for the process of first language acquisition, it is impossible

to pinpoint when one has finally acquired the entire system, for people are constantly learning new words, new (idiomatic, technical, slang) expressions and new pragmatic uses of language (Singleton & Ryan, 2004). With L2 development, the ongoing change of the system is even more evident, for learners will always have room for improvement in their fluency, accuracy, proficiency, competence, intelligibility, comprehensibility, etc.

As mentioned above, the prototypical phonological categories of the L1 may act as attractor states for the L2 developing system, causing Brazilian learners to perceive (and thus produce) only [i] when exposed to [i] or [ɪ], for example. Therefore, one of the purposes of the language classroom is to help learners move their systems away from these attractor states into the ones that have the appropriate distinctions between L2 contrasting sounds. Some learners need more intervention to have their systems exit an attractor state than others, but they can all potentially do so.

This article hypothesizes that for some Brazilian undergraduate learners of English, the English Phonology course may function as an intervention strong enough to help them move their developing L2 vowel systems into a state with the appropriate contrasting categories. It is also expected, though, that the course will not be enough for some learners to create new vowel categories for the L2, at least not immediately. Because of the non-linear relation between perturbation (intervention – phonology classes) and movement of the system (creation of new vowel categories), it is possible that the effect of the lessons will appear later on for some learners.

This dynamic nature of L2 phonological development is what makes dynamic systems better examined in a longitudinal study rather than in a cross-sectional one (De Bot & Larsen-Freeman, 2011; Larsen-Freeman & Cameron, 2008; Lima Jr, 2016a; Verspoor, De Bot, & Lowie, 2011). Also, in a dynamic system, the processes are more relevant than the products, since its ongoing change in time makes it impossible for the system to reach an end state. Hence, concepts such as *final state* or *ultimate attainment* should be replaced by concepts of *dynamic idiosyncratic development*.

It is for those reasons that the data presented in this article come from an umbrella longitudinal project which looks into the phonological development of English-L2 learners individually. The participants, 13 Brazilian undergraduate learners of English Language Teaching (ELT), have been recorded every semester since they were admitted to college, and will keep being recorded every semester until they graduate. The data presented here zoom in at the vowel production of the second and third recordings, collected right before and right after students took the mandatory English Segmental Phonology course in their third term.

The main goal of the umbrella project is to investigate the individual routes of English phonological development by these Brazilian undergraduate students. Within this general goal, this article has the more specific objective of analyzing the individual creation of new vocalic categories, in terms of spectral quality (F1-F2), after having received explicit metalinguistic instruction on English segmental phonology. As a result, the main goal of this article is to discuss the influence of metalinguistic knowledge of segmental phonology on the production of English vowels by Brazilian undergraduate students of English.

2 Method

The data come from 13 Brazilian undergraduate students majoring in English Language Teaching (ELT) at a federal university in the state of Ceará, Brazil. As part of the umbrella longitudinal project under which this study lies, the students started being recorded every semester since their admission to college and will be recorded every semester until they graduate. This article presents data from the second and third recordings, which were done before and after their third term, when they take a mandatory course on English Segmental Phonology.

The phonology course is 64 hours long, lasting one entire school semester, and is taught in English. It is a technical and metalinguistic course, but with a secondary goal of helping the non-native English speaking teacher students to improve their pronunciation of English

vowels and consonants. The course begins with the basics of articulatory phonetics (8 hours) and the principles of the International Phonetic Alphabet – IPA (4 hours). The remainder is equally divided into the study of English consonants and English vowels. Students learn how to transcribe words using the IPA and to read transcribed words. They learn how to classify consonants into place of articulation, manner of articulation and voicing; and the vowels into tongue height, tongue advancement, lip position and muscle tension (tense vs lax). They also study the relations between orthography and phonology, and, towards the end of the course, reflect on how to teach the sounds of English consonants and vowels to Brazilians.

The undergraduate students were recorded individually, in a silent room, reading words inserted in the carrier sentence “I said *token* this time”. The corpus was composed of three words for each target vowel. The words were controlled for phonological context – all of them were monosyllabic with a CVC context, where both Cs were voiceless plosives.² This control was meant to prevent acoustic bias from the neighboring segments, and it also made it easier to identify, segment and label the target vowels in PRAAT (Boersma & Weenink, 2011), the software used to conduct the acoustic analyses. The words are presented in Table 1.

Table 1: Corpus for data collection for acoustic analysis

[i]	[ɪ]	[ε]	[æ]	[u]	[ʊ]
peak	pick	peck	pack	boot	book
Pete	Pitt	pet	pat	poop	put
teak	tick	tech	tack	toot	took

The words were presented to participants in the carrier sentence “I said *token* this time” (Watkins & Rauber, 2010), which controls for the number of syllables before and after the target word, preventing, thus, intonational bias from the beginning and end of sentences read as off a list. The sentences were presented in a slides presentation, with each sentence on an individual slide. Each word was randomly repeated four times, generating 12 tokens per vowel per participant,

which generated 72 tokens per participant, and a total of 936 vowels per semester. In the end, a total of 1,872 vowels were identified and analyzed.

The recordings were done with a supercardioid Shure 150B lapel microphone connected to a Zoom 4HnSP recorder. The audio was captured in mono, with a sampling rate of 44 KHz, and saved in wav format.

The vowels were segmented in PRAAT (Boersma & Weenink, 2011). The points considered as beginning and end of each vowel were the first and last valley in the periodic pulse in the waveform which had considerable amplitude, resembled the vocalic period, and presented stable formants in the spectrogram.

One of the most common methods used to extract formants is the Linear Predictive Coding (LPC), which is a predictive algorithm that decomposes the acoustic signal and makes an estimation of the resonances generated in the vocal tract. However, the automatic LPC analysis has been criticized (e. g. Vallabha & Tuller, 2002; Wempe & Boersma, 2003) because it may introduce systematic errors in the formant extraction depending on the parameters set beforehand by the researcher. The problem is that, with the automatic LPC analysis, the researcher needs to define, before the analysis, the order of the LPC (i.e., the quantity of formants to be found) and the maximum (ceiling) frequency in which to look, which is usually set as 5 KHz for men and 5.5 KHz for women. Nevertheless, different men and women might have different frequency ceilings, which, if not set accordingly, might lead the LPC into identifying peaks that do not exist and overlooking peaks that do.

A solution to this problem is to double-check the adjustment of the LPC to the FFT spectrum (obtained by the Fast Fourier Transform algorithm) vowel by vowel. Even though this method is more time-consuming, it allows the researcher to adjust, when necessary, the ceiling frequency or the order of the LPC for specific speakers. This is what the scripts used to extract F1 and F2 in this study do (Arantes, 2010, 2011).

After extracted, the F1 and F2 values were used to create vowel space plots using the package PhonR (McCloy, 2016) for the software R (R Core Team, 2016). The same package was used to normalize the formant

values using the Lobanov method, which creates a z-score for F1 and F2. This was done in order to later calculate the Euclidean Distances between the vowels without the bias of F2 values, which have raw values that are much larger and that increase in much larger increments than F1. Finally, t-tests were conducted with F1 and F2 values of target pairs of vowels.

3 Results

The first step in the data analysis was to visually inspect the individual vowel spaces, comparing the distributions of speakers' vowels in the second and third recordings. Since the plots present, for each vowel, all the tokens produced by the speaker, the point of intersection between the F1 and F2 means and an ellipsis with one standard deviation, it was relatively easy to visually identify vowels that overlapped and vowels that were produced separately.

In this first visual inspection of plots, when two vowels had half or more of their one-standard-deviation ellipses overlapping, they were considered overlapping vowels (i.e., only one vocalic category for both); when less than half of the ellipses overlapped or when they did not overlap at all, they were considered separate vowel categories. As will be shown later, two other more quantitative methods were used to ratify this somewhat qualitative classification.

As an example, the image below features the plots of two learners in the second recording (before the phonology classes), one of whom (speaker A) had two separate categories for the vowels [i ɪ], but had overlapping vowels for the other two pairs; and the other student (speaker D) had overlapping vowels for all three pairs. To avoid the risk of losing information when using phonetic symbols in different computer programs (excel, notepad, PRAAT and R), Well's (1982) keywords for the English language were used instead. Therefore, where there is FLEECE, KIT, DRESS, TRAP, GOOSE and FOOT in the plots throughout the article, please read [i ɪ ε æ u ʊ], respectively.

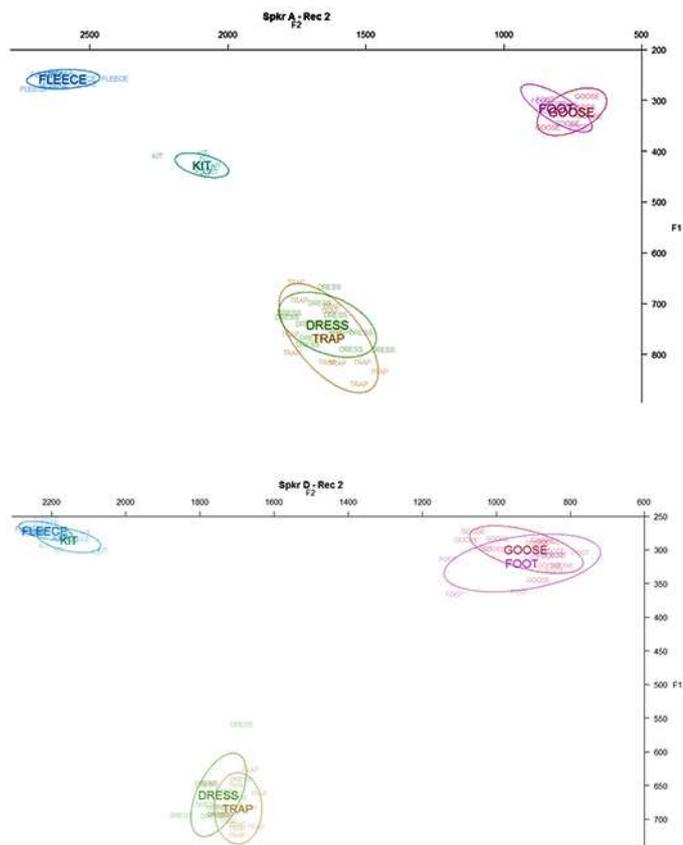


Figure 1: Vowels spaces with and without contrast for the vowels [i ɪ]

As can be seen, speaker A had separate vowel categories for [i ɪ] before he took the phonology course. Besides speaker A, other 7 speakers (A, B, F, G, K, L and N) had separate categories for the high front vowels before taking the phonology course. In relation to the pair [u ʊ], three learners (G, I and J) had separate vowel categories before the phonology course. Notice that speaker G is also listed as having the [i ɪ] contrast in this recording, making him the only learner with two contrasts before the phonology course. His vowel space can be seen in Figure 2.

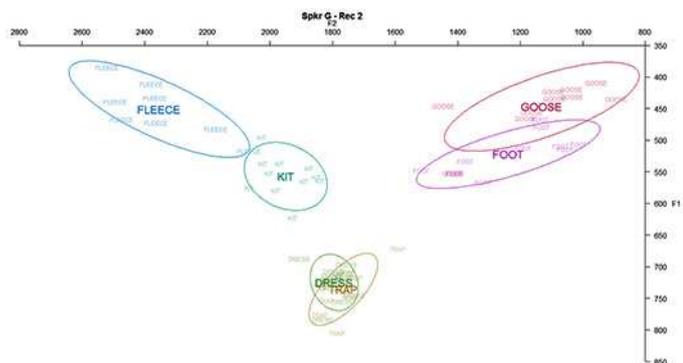


Figure 2: Speaker's G vowels space before the phonology course

In the recording done before the phonology course, no distinction was produced in the [ɛ æ] contrast. The results presented up to this point already reveal a hierarchy of difficulty for Brazilian learners of English concerning these three pairs of vowels, with [ɛ æ] being the hardest and [i ɪ] the least difficult. This result is similar to those found in Lima Jr (2015), Barboza (2008) and Rauber (2006).

After the course on English Phonology, the learners were recorded again and their vowel spaces were created and visually analyzed in the same manner, but this time comparing their productions before and after the lessons in order to look into the effects of the explicit instruction. Besides the individual vowels spaces, a third plot was created for each speaker, with the tokens and the mean points for all vowels color-coded by recording. An example of this plot, from speaker A's productions, can be seen in Figure 3, along with the corresponding separate vowel spaces.

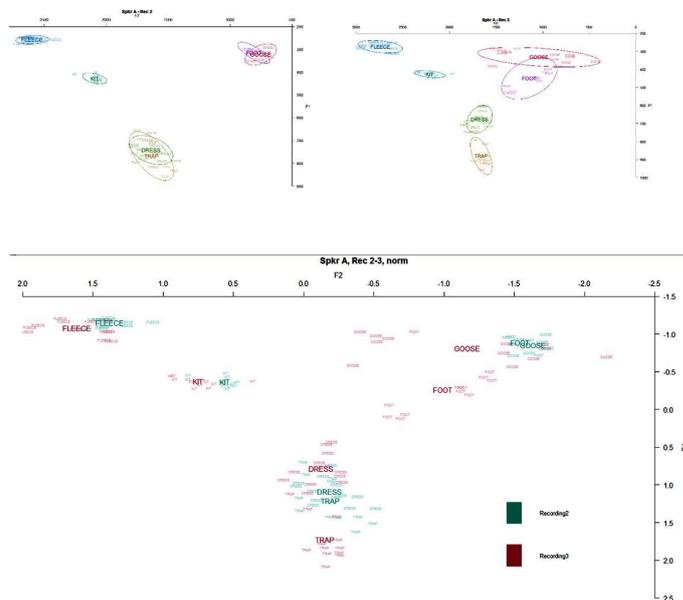


Figure 3: Plots comparing speaker A's recordings 2 and 3 (before and after the phonology course)

As can be seen, this new plot with data from the two recordings contains Lobanov-normalized vowels. This was done in order to decrease possible effects from the different elocutions the same speaker might have in recordings done with a 6-month interval, as well as to have both F1 and F2 in the same scale for further calculation of Euclidean Distances without the bias

F2 values, which inherently increase in a much higher increment than F1.

It can also be seen in Figure 3 that this particular learner (speaker A) created new categories for both [u ʊ] and [ɛ æ] while taking the course on segmental phonology. Since he already had the [i ɪ] contrast before the course, he ended the semester with contrasts in all three pairs. The third plot clearly shows how close his FOOT-GOOSE and DRESS-TRAP means were in recording 2 (before the phonology course) and how distant they became in recording 3 (after the course).

Besides speaker A, two more speakers created new categories for [u ʊ] (E and N), and two more were now able to produce the [ɛ æ] contrast (F and N). Also, two learners created new [i ɪ] categories (I and J). Table 2 presents a summary of the presence of vowel contrasts in recording 2 (before the phonology course) and the creation of new vowel contrasts in recording 3 (after the lessons on phonology).

Table 2: Vowel contrasts in recordings 2 and 3, before and after the phonology course

Speaker ³	Recording	Contrast?		
		[i ɪ]	[ɛ æ]	[u ʊ]
A	2	YES	no	no
	3	YES	YES	YES
B	2	YES	no	no
	3	YES	no	no
D	2	no	no	no
	3	no	no	no
E	2	no	no	no
	3	no	no	YES
F	2	YES	no	no
	3	YES	YES	no
G	2	YES	no	YES
	3	YES	no	YES
I	2	no	no	YES
	3	YES	no	no
J	2	no	no	YES
	3	YES	no	YES
K	2	YES	no	no
	3	YES	no	no

L	2	YES	no	no
	3	YES	no	no
M	2	no	no	no
	3	no	no	no
N	2	YES	no	no
	3	YES	YES	YES
O	2	no	no	no
	3	no	no	no

The presence of a contrast is signaled with the word YES, and the creation of new categories in the third recording, besides having the word YES, is also highlighted. There was a total of 10 contrasts in recording 2, and other 8 were created after the phonology course. This allowed for two learners (A and N) to end their third term with six well-defined categories for these six English vowels that are challenging for Brazilians. In the third recording, the hierarchy of difficulty found in the previous recording was maintained, with 9 students producing the [i ɪ] contrast, 5 producing the [u ʊ] contrast, and 3 producing [ɛ æ] separately.

On the negative side, 8 students made absolutely no progress from one recording to another, of which a total of 5 (D, E, I, M and O) had all their pairs of vowels overlapping in both recordings. Also, one learner (speaker I) made a [u ʊ] contrast in the second recording (before the phonology lessons), which disappeared in the third recording.

Two mathematical instruments were used in order to ratify the conclusions reached so far. The first one was the measurement of the Euclidean Distances between the mean points of contrasting vowels for each speaker. The Euclidean Distance is a measure of dissimilarity used to compare two or more items given a number of quantifiable characteristics. When the number of quantifiable characteristics is two, it can be used to measure the distance between two points in a cartesian coordinate system, which is the case of the F1-F2 graph. Its formula is derived from the Pythagorean theorem, and, taking the [ɛ æ] contrast as an example, it was calculated as:

$$Eucl. Dist. = \sqrt{(F1\ of\ [\varepsilon] - F1\ of\ [\varepsilon])^2 + (F2\ of\ [\varepsilon] - F2\ of\ [\varepsilon])^2}$$

As mentioned in the method section, the formant values for each speaker were previously Lobanov-normalized

so that the Euclidean Distances did not include the bias of the F2 values, which are inherently larger and increase in larger increments than those of F1.

Figure 4 presents speaker A's vowel spaces for both recordings, with a visual representation of the concept of Euclidean Distances on a cartesian coordinate system. The idea was to see if the distances between the contrasting vowels [i ɪ], [u ʊ] and [ɛ æ] increase from one recording to the next.

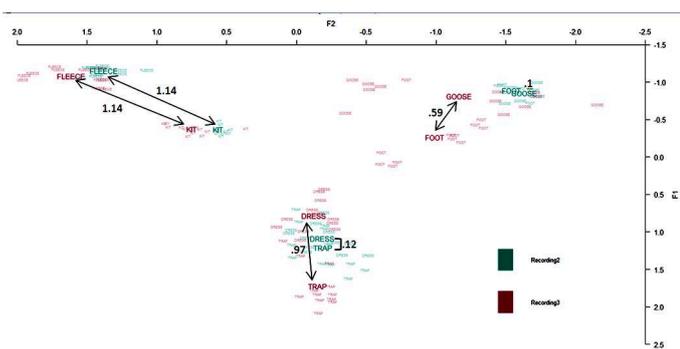


Figure 4: Speaker A's vowel spaces with Euclidean Distances for contrasting vowels

As can be seen, speaker A maintained the exact same distance between [i ɪ], which was his largest distance, in both recordings (1.14). His [u ʊ] distance increased from 0.12 to 0.97 in the recording after the phonology lessons, and his [ɛ æ] distance increased from .0.10 to 0.59.

The following table presents the Euclidean Distances for all vowel contrasts for all speakers in both recordings.

Table 3: Euclidean Distances for vowel contrasts in both recordings

Speaker	Recording	Vowel Contrast		
		[i ɪ]	[ɛ æ]	[u ʊ]
A	2	1.14	.12	.10
	3	1.14	.97	.59
B	2	1.31	.12	.13
	3	1.19	.05	.34
D	2	.15	.15	.22
	3	.26	.07	.16
E	2	.14	.15	.52
	3	.33	.12	.90
F	2	.6	.21	.37
	3	.53	.65	.14

G	2	1.31	.16	.60
	3	1.68	.28	.87
I	2	.25	.23	.65
	3	.55	.12	.26
J	2	.31	.25	.58
	3	.78	.11	.89
K	2	.71	.27	.10
	3	.74	.39	.33
L	2	.93	.03	.47
	3	.95	.35	.64
M	2	.24	.45	.24
	3	.08	.68	.08
N	2	1.47	.09	.41
	3	1.38	.49	.87
O	2	.04	.11	.36
	3	.41	.16	.09

In the preceding table, all vowel contrasts present in recording 2 and kept in recording 3 were highlighted in yellow, and all the contrasts created only in recording 3 were highlighted in red. Notice that all contrasts, either present and kept or newly created, have Euclidean Distances of at least 0.5. Also, in the contrasts created in recording 3, the difference between the Euclidean Distances is of at least .30.

In Table 3 there are very few distances higher than .5 that were not considered contrastive because, even though the mean points are slightly apart, the standard deviations are very high, causing the ellipses to overlap, which shows that there is no phonological contrast. For instance, this was the case of speaker L's productions of [u ʊ] in recording 3, whose vowel space is presented in Figure 5. Notice that FOOT and GOOSE are somewhat apart, but with overlapping ellipses.

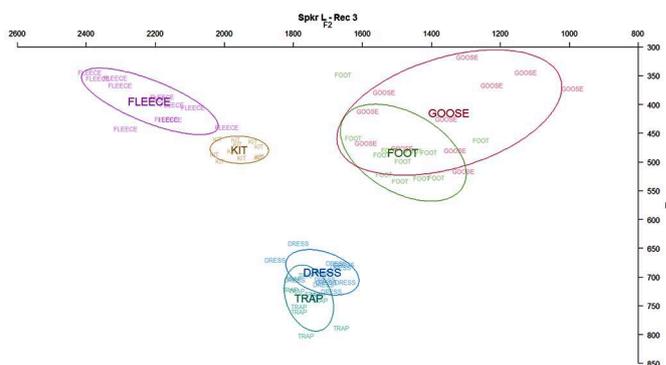


Figure 5: Speaker L's vowel space in recording 3, illustrating a slight distance between [u] and [ʊ], but with large ellipses (one standard deviation) that overlap

The second mathematical instrument used to verify the conclusions of the visual inspection of plots was a statistical significance test, specifically t-tests with formant values of target pairs of vowels in each recording. To reduce the size of the table and increase readability of the results, only the p-values of the t-tests are presented in the following table, and only of the t-tests conducted with F1 values. Since the three pairs of vowels in focus contrast mainly in tongue height, the results of the t-tests with F1 values are more informative than those with F2 values.

Table 4: P-values for t-tests conducted with F1 values of contrasting pairs of vowels

Speaker	Recording	T-Tests for F1 (p values)		
		[i ɪ]	[ɛ æ]	[u ʊ]
A	2	.00	.24	.62
	3	.00	.00	.00
B	2	.00	.40	.07
	3	.00	.64	.01
D	2	.00	.16	.06
	3	.12	.63	.00
E	2	.24	.92	.02
	3	.00	.27	.00
F	2	.00	.00	.00
	3	.89	.00	.06
G	2	.00	.38	.00
	3	.00	.00	.00
I	2	.19	.27	.00
	3	.00	.57	.03
J	2	.10	.01	.02
	3	.00	.28	.00
K	2	.00	.02	.30
	3	.03	.00	.20
L	2	.00	.92	.03
	3	.00	.00	.00
M	2	.20	.10	.14
	3	.43	.00	.44
N	2	.00	.24	.02
	3	.00	.00	.00
O	2	.51	.32	.08
	3	.06	.03	.23

The results from the t-tests reveal that all the pairs of vowels considered contrastive with the visual inspection of the vowel spaces and the calculation of Euclidean Distances had a significant p-value (with alpha at 5%), showing that the two contrasting vowels were also statistically different. The only exception was with the [i ɪ] contrast of speaker F in recording 3, whose p value is 0.89. This was the only case in which the t-test conducted with F2 values was useful to explain a contrasting pair of vowels. Looking at this speaker's vowel plot (Figure 6), it is clear that, in the third recording, he separated the two vowels much more in terms of tongue advancement than tongue height. The p-value of the t-test conducted with the F2 values was 0.00.

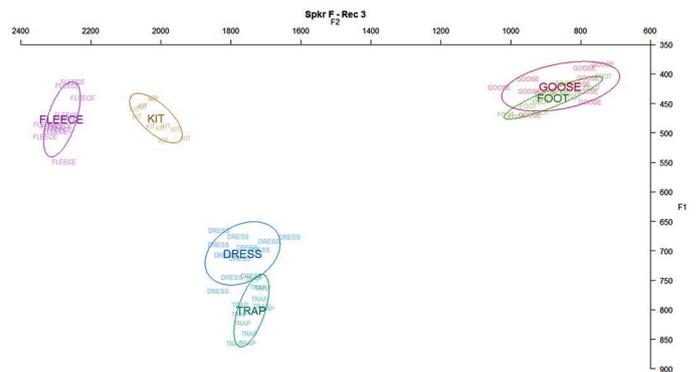


Figure 6: Speaker F's vowel space in the third recording, showing that his [i ɪ] contrast was much higher in F2 and F1

Referring back to the results of the t-tests in Table 4, one can notice that there are a few significant p-values for pairs of vowels that were not considered contrastive. This was the same case with the Euclidean Distances higher than 0.5 not considered contrastive; that is, it happened when the means of target vowels were somewhat separate, and picked by the statistical test, but the standard deviations were so high that the ellipses in the vowel spaces overlapped at least 50%. An example of such situation is the production of [i ɪ] in recording 2 and of [u ʊ] in recording 3 by speaker D, both with p-values of 0.00 yet clearly overlapped in the vowel spaces (Figure 7).

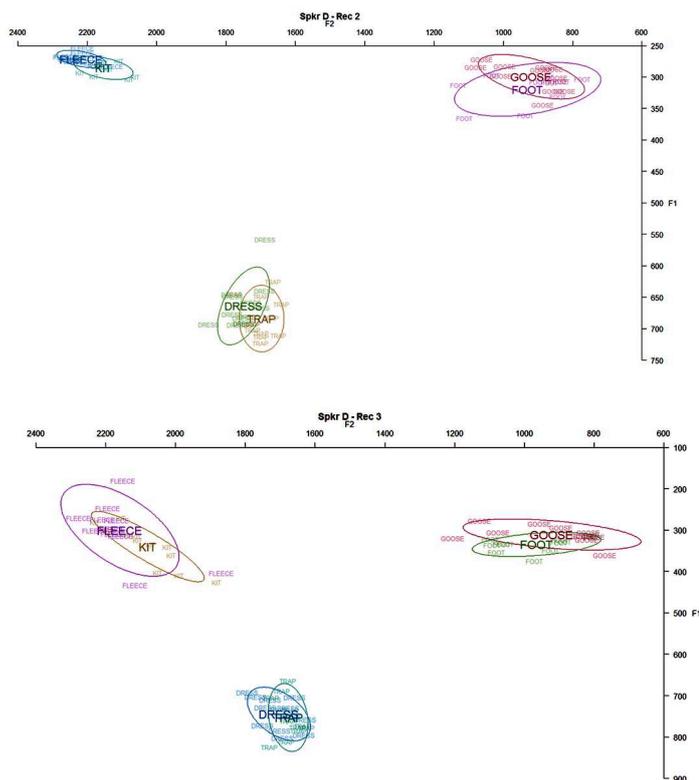


Figure 7: Speaker D's vowels spaces for recordings 2 and 3, to illustrate that t-tests with formant values not always correspond to distinctive vowels

Using the information from the visual inspection of individual vowel spaces together with the Euclidean Distances and the results from the t-tests, it was possible to verify the contrasts present in recording 2 (before the phonology course) and kept in recording 3 (after the phonology course) as well as those created only in recording 3, possibly due to the effects of receiving explicit metalinguistic instruction on English Segmental Phonology, as will be further discussed in the following section.

4 Discussion

The results first showed that a few learners already had some vowel contrasts before taking the course on English Segmental Phonology. This was expected for two main reasons. First, Brazilian undergraduates of English Language Teaching (ELT) are admitted to college without taking an English proficiency test. This means that their proficiency levels vary a lot, especially in initial terms, which accounts for the

presence of some vowel contrasts in the speech of the more advanced learners. However, not even the more advanced learners produced the contrasts for all three pairs of vowels. Before taking the phonology course, from the 13 students recorded, 8 had contrasting pairs of vowels, of which only one had contrasts for two pairs.

The second reason for such expectation is that there is a lot of individual variability among learners of English in a non-native English speaking country, such as Brazil. English is a mandatory subject in the four years of Middle School⁴ and in the three years of High School.⁵ However, English instruction in regular school emphasizes writing skills, and students develop different levels of interest in developing their oral skills on their own. Also, many private schools in Brazil offer English classes in the five years of Elementary School,⁶ and many teenagers take English classes in extracurricular language institutes, where there is usually more emphasis on oral communication. All these differences in amount of instruction before college education would already account for such variability, not to mention that different teenagers get exposed to English in different amounts through media (music, movies, TV series), vlogs, YouTube channels, video games, computer programs, podcasts, etc.

From a Dynamic Systems Theory (DST) perspective, the fact that each student in class is at a different developmental stage is taken for granted. Even if they take placement tests before starting a course, different students will never be at the exact same point in L2 development. Teachers need to acknowledge, and keep reminding themselves, that each language student is a dynamic system undergoing a process of language learning which is also a dynamic system. Each system is made up of so many elements, whose interaction among themselves and with the environment make the performance in the L2 emerge, that it is impossible to expect all students to be at the same initial stage in the beginning of a course.

In addition, due to the non-linear relation between cause and effect in a dynamic system, language instructors should not expect learners to react the same way to teaching interventions. In the case of this study, it was expected that each student would react

differently to the metalinguistic instruction on English Segmental Phonology. The same way a lot of snow might just accumulate on a mountain top, not causing any immediate effect, and sometimes the slightest movement on a mountain, such as the presence of a skier, might cause an avalanche, the classroom interventions might not cause any movement in some L2 learners' systems, at least not immediately, while they might cause a lot of movement in other learners' L2 developing system.

The creation of new vowel categories after the phonology course confirms such expectation. In recording 3, six learners had created new vowel contrasts, and two of them ended up having contrasting vowels for all three target pairs. This does not mean, though, that the lessons did not have any effect on the other learners; it simply means that no effect was found at that moment. It is possible that in later stages of development, and triggered by other perturbations of their systems, those learners that showed no immediate effect will easily move their systems away from the attractor states of the prototypical vowel categories of their L1. The overall progress of the class, however, was satisfactory. This is especially true if compared to the changes from recording 1 to recording 2, reported in Lima Jr (2016b), when only two learners created new categories for one pair of vowels each (speakers G and J for the contrast [u ʊ]).

The learner that apparently “unlearned” a vowel contrast after the phonology course is also contemplated by DST. Speaker I produced [u] and [ʊ] separately in the recording before the phonology course, but overlapped in the next recording. This is an instance of the ups and downs, peaks and valleys, through which learners go in the process of learning a new language. The contrast produced in recording 2 was probably not yet stabilized, and thus was not present in recording 3. That learner has probably already understood that there is something different in the production of these high back vowels, but still needs to develop consistency in his production.

The results also reinforce the order of difficulty that those three pairs of vowels pose to Brazilian learners of English. As mentioned in the results section, the most

difficult pair for Brazilians is [ɛ æ], followed by [u ʊ], with the least difficult being [i ɪ]. It is true that native speakers of some dialects of English do not produce [ɛ æ] with spectral difference, for instance speakers of the Northern Cities Dialect of the United States (Ladefoged & Johnson, 2011), in the areas around Detroit and Rochester. Those speakers contrast this pair of vowels by lengthening the front open one, [æ]. Not having looked into duration yet is a limitation of this study, which will be easily remedied in a later stage of the project, since all vowels have already been segmented and labeled in PRAAT.

A final topic of discussion, which did not come up in the results section, has to do with the gradient view of language development under a DST perspective. The results section was an attempt to categorize students' productions into “contrasting vowels” and “no contrasting vowels”. However, language development and language production are not categorical, but rather gradient in nature. This means that it was not always easy to decide if two vowels should be considered “with” or “without” a contrast. That is why some criteria needed to be defined (e.g. overlap of at least half of the ellipses, Euclidean Distances higher than .5, etc.) and followed for the categorization of the results.

Nevertheless, under a DST perspective, one cannot overlook the gradience found in the data. Some students were classified into “no contrast”, but were almost creating new categories. The binary classification of participants may give the wrong impression that all learners with a “no” in Table 2 produced the contrasts equally overlapped, which was not the case. Some students have moved their vowels apart, yet not enough as to fulfill the classification criteria established. This was the case with the production of [ɛ æ] by speaker K. His vowel spaces for recordings 2 and 3 (Figure 8) show that his [ɛ] and [æ] are slightly more separated in the latter, but just not yet as to fulfill the criteria.

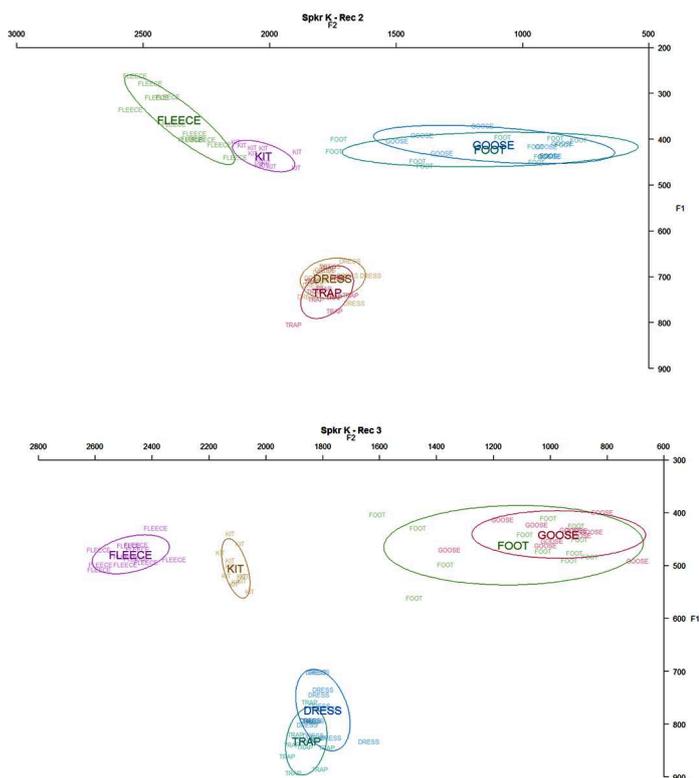


Figure 8: Speaker K's vowel spaces for recordings 2 and 3, showing a slight, yet not satisfactory movement in the [æ] [ɛ] contrast

Likewise, one might think that all speakers registered with contrasting vowels in the results section produced the contrasts equally well. However, some learners produced contrasting vowels in the threshold of the criteria established, whereas others produced vowels truly separated, with the ellipses far from touching each other. As a matter of fact, the data present variation even within the same speaker. Speakers F, G and K, for instance, are all marked with separate categories for [i] and [ɪ] in the second and third recordings. While this is true, this simple categorization overlooks the fact that their contrasts in the third recording were much higher than in the previous one. This is shown in Figure 9, comparing the [i] [ɪ] contrasts of speaker G, as an example.

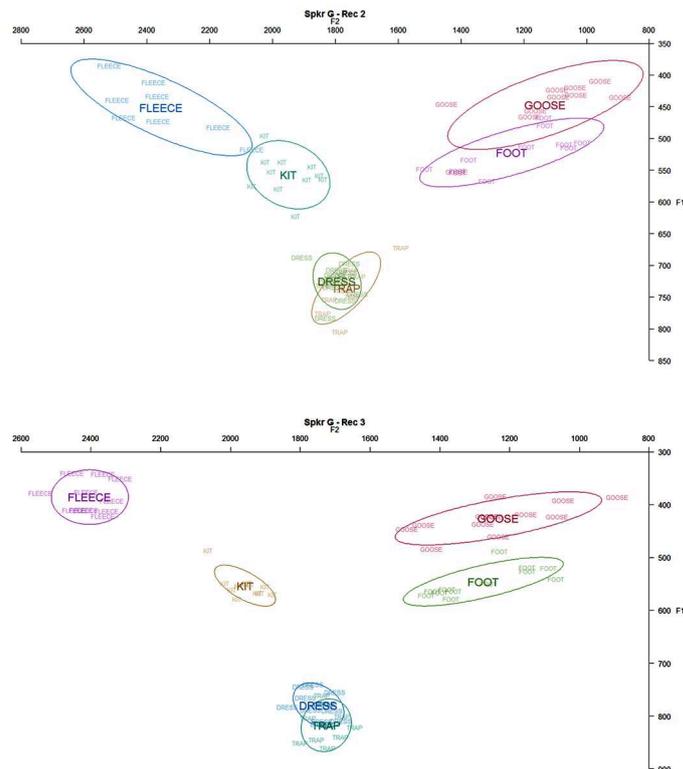


Figure 9: Speaker G's vowel spaces for recordings 2 and 3, showing that the latter has a greater distance between [i] and [ɪ]

As can be seen, even though he has [i] [ɪ] contrasts in both recordings, the one in the latter is much clearer, showing that he has made progress with vowel categories that were already somewhat resolved. A data analysis that looks exclusively at group means and/or looks exclusively at classifications is unable to register and account for the nuances of the individual routes of L2 development, which can be very informative concerning the process of language acquisition.

5 Conclusion

The main goal of this article was to discuss the influence of explicit metalinguistic instruction of segmental phonology on the production of six English vowels by Brazilian undergraduate students of English Language Teaching. This goal was achieved by analyzing the individual creation of new vocalic categories, in terms of spectral quality (F1-F2), after having taken an undergraduate-level course on English Segmental Phonology. With the recording done after

the phonology course, it was possible to see that six students had created new vowel categories, and two of them created new categories for more than one pair of target vowels. Therefore, the conclusion is that a metalinguistic course on English phonology has positive effects on the production of vowels, and hopefully on the overall pronunciation, of some learners. Colleges should, hence, keep offering such a course in ELT undergraduate programs.

This study has limitations, most of which can be resolved in future stages of the umbrella longitudinal project. The first one, which has already been mentioned, regards the inclusion of a vowel duration analysis. This is especially important because all the three pairs of vowels are usually also contrasted in duration by native speakers of English. This measure will be taken in a future analysis.

The second limitation regards the extremely controlled data collection, with participants reading words in a carrier sentence. Even though this design is common practice is laboratory phonology, it moves away from a more realistic and authentic use of language. All participants in this project, though, have also been recorded reading a paragraph, whose vowels will be analyzed in a future stage, and results will be compared with those herein reported. The hypothesis is that more vowel contrasts will be present in the controlled design (reading carrier sentences), when the participants have more focused attention to their articulation and, therefore, can monitor their pronunciation more carefully.

The last limitation worth mentioning is the need to include a perception experiment, which would allow for the comparison of perception and production of English vowels through time, hopefully witnessing the emergence of both perceptual and productive vowel categories. Unfortunately, these data have not been collected from the participants of this study, leaving this final limitation as a suggestion and recommendation for future similar studies.

Acknowledgements

This project has been partially funded by the Brazilian National Council for Scientific and Technological Development (CNPq), grant number 471868/2014-0.

Notes

1. This is why this article prefers the term “development” over “acquisition”, and only uses the term “acquisition” if preceded by “the process of (acquisition)”.
2. Speakers C and H are missing because they did not participate in the third recording. Losing participants is an unfortunate challenge in longitudinal studies.
3. With the exception of two words with the back vowels, due to the lack of words with such features.
4. Ensino Fundamental II
5. Ensino Médio
6. Ensino Fundamental I

References

- Arantes, P. (2010). Formants.praat, v. 0.9 beta.
- Arantes, P. (2011). Collectformants.praat, v. 0.11 alpha.
- Barboza, C. L. F. (2008). *Descrição acústica dos sons vocálicos anteriores do inglês e do português realizados por professores de inglês língua estrangeira no oeste potiguar. 183 f.* Dissertação (Mestrado Acadêmico em Linguística Aplicada)–Universidade Estadual do Ceará, Fortaleza.
- Bion, R. A. H., Escudero, P., Rauber, A. S., & Baptista, B. O. (2006). Category formation and the role of spectral quality in the perception and production of English front vowels. In *Interspeech*. Citeseer, 1363-1366.
- Boersma, P. & Weenink, D. (2011). Praat: doing phonetics by computer [Computer program]. Version 6.0.30, retrieved from <http://www.praat.org/>.
- Bybee, J. (2003). *Phonology and language use* (Vol. 94). Cambridge: Cambridge University Press.
- Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (2010). *Teaching Pronunciation: A Course Book and Reference Guide* (2nd ed.). Cambridge: Cambridge University Press.
- Cristófaro Silva, T. (2003). Descartando fonemas: a representação mental da fonologia de uso. In D. Hora & G. Collischonn (Eds.), *Teoria linguística: fonologia e outros temas*. João Pessoa: Editora Universitária.

- De Bot, K. (2008). Introduction: Second language development as a dynamic process. *The Modern Language Journal*, 92(2), 166–178.
- De Bot, K., & Larsen-Freeman, D. (2011). Researching second language development from a dynamic systems theory perspective. *A Dynamic Approach to Second Language Development: Methods and Techniques*, 5–24.
- De Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism Language and Cognition*, 10(1), 7.
- Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). Baltimore: York Press.
- Flege, J. E. (1999). Age of learning and second language speech. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 101–131). Mahwah: Lawrence Erlbaum Associates.
- Flege, J. E. (2007). Language contact in bilingualism: Phonetic system interactions. *Laboratory Phonology*, 9, 353–382.
- Kuhl, P. (1991). Human adults and human infants show a “perceptual magnet effect” for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics*, 50(2), 93–107.
- Kuhl, P. K. (1993). Innate predispositions and the effects of experience in speech perception: The native language magnet theory. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. McNeilage, & J. Morton (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 259–274). The Hague: Kluwer Academic Publishers.
- Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363(1493), 979–1000.
- Ladefoged, P., & Johnson, K. (2011). *A course in phonetics* (6th ed.). Boston, MA: Wadsworth.
- Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics*, 18(2), 141–165.
- Larsen-Freeman, D., & Cameron, L. (2008). Research methodology on language development from a complex systems perspective. *The Modern Language Journal*, 92 (2), 200–213.
- Leather, J. (2003). Phonological acquisition in multilingualism. In M. del P. G. Mayo & M. L. G. Lecumberri (Eds.), *Age and the Acquisition of English as a Foreign Language* (pp. 23–58). Clevedon: Multilingual Matters.
- Lima Jr, R. M. (2013). Complexity in Second Language Phonology Acquisition. *Revista Brasileira de Linguística Aplicada*, 13(2).
- _____. (2015). A influência da idade na aquisição de seis vogais do inglês por alunos brasileiros. *Organon*, 30(58), 17.
- _____. (2016a). A necessidade de dados individuais e longitudinais para análise do desenvolvimento fonológico de l2 como sistema complexo. *ReVEL*, 14(27), 203–225.
- _____. (2016b). Análise longitudinal de vogais do inglês-L2 de brasileiros: dados preliminares. *Gradus: Revista Brasileira de Fonologia de Laboratório*, 1(1), 145–176.
- McCloy, D. R. (2016). phonR: tools for phoneticians and phonologists.
- Morley, J. (1994). *Pronunciation pedagogy and theory: new views, new directions*. Bloomington: ERIC.
- Nobre-Oliveira, D. (2007). *The effect of perceptual training on the learning of English vowels by Brazilian Portuguese speakers*. Unpublished doctoral dissertation, Federal University of Santa Catarina, Florianópolis, Brazil.
- Parrino, A. (1998). The politics of pronunciation and the adult learner. *Adult ESL: Politics, Pedagogy and Participation in Classroom and Community Programs*, 171–184.
- Pennington, M. C. (1996). *Phonology in English Language Teaching*. New York: Addison Wesley Longman Limited.
- _____. (1998). The Teachability of Phonology in Adulthood: A Re-Examination. *IRAL*, 36(4), 323–341.
- Pierrehumbert, J. (2001). Stochastic phonology. *Glott International*, 5(6), 195–207.
- _____. (2003). Phonetic diversity, statistical learning, and acquisition of phonology. *Language and Speech*, 46(2–3), 115–154.
- R Core Team. (2016). R: A language and environment for statistical computing.
- Rauber, A. S. (2006). *Perception and production of English vowels by Brazilian EFL speakers*. Universidade Federal de Santa Catarina, Florianópolis.
- Silveira, R. (2004). *The influence of pronunciation instruction on the perception and the production of English word-final consonants*. Universidade Federal de Santa Catarina.

- Singleton, D. M., & Ryan, L. (2004). *Language acquisition: the age factor* (Vol. 47). Clevedon, New York, Ontario: Multilingual Matters.
- Tomlinson, B. (2005). English as a foreign language: Matching procedures to the context of learning. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (Vol. 1, pp. 137–154). Mahwah: Lawrence Erlbaum Associates.
- Vallabha, G. K., & Tuller, B. (2002). Systematic errors in the formant analysis of steady-state vowels. *Speech Communication*, 38(1), 141–160.
- Verspoor, M., De Bot, K., & Lowie, W. (2011). *A dynamic approach to second language development: Methods and techniques* (Vol. 29). Amsterdam: John Benjamins Publishing.
- Watkins, M. A., & Rauber, A. S. (2010). Variability in pretonic vowel reduction by Brazilian speakers of English. In A. S. Rauber, M. A. Watkins, R. Silveira, & R. Koerich (Eds.), *The acquisition of language speech: studies in honor of Professor Barbara O. Baptista* (pp. 75–99). Florianópolis: Insular.
- Wells, J. C. (1982). *Accents of English* (Vol. 1). Cambridge: Cambridge University Press.
- Wempe, T., & Boersma, P. (2003). The interactive design of an F0-related spectral analyser. In *Proc. 15th ICPhS* (pp. 343–346).
- Yule, G., & Macdonald, D. (1994). The effects of pronunciation teaching. *Pronunciation Pedagogy and Theory: New Views, New Directions*, 109–118.

Recebido em: 31/03/2017

Aceito em: 26/07/2017