A LONGITUDINAL STUDY ON THE ACQUISITION OF SIX ENGLISH VOWELS BY BRAZILIAN LEARNERS

Ronaldo Lima Jr

Federal University of Ceará
ronaldojr@letras.ufc.br

ABSTRACT

This study analyzed the production of English vowels [i ɪ ɛ æ u ʊ] by Brazilian undergraduate students of English Language Teaching throughout the four first semesters of their studies. These vowels are challenging for Brazilians for they are not contrasted in Brazilian Portuguese. Ten participants were recorded reading target words in a carrier sentence every semester from semester 1 through 4. Euclidean distances between pairs of vowels were calculated using Lobanov-normalized F1-F2 values, and such distances were used to fit a mixed-effects model to the data. The results show that: most learners increased their contrasts of the target vowels; learners developed their vowels at different paces and in different moments; not all learners were able to create new phonetic categories for the target vowels.

Keywords: English as a foreign language, vowels, L2 acquisition, longitudinal

1. INTRODUCTION

The main goal of this study was to analyze the development of English vowels [i ɪ ɛ æ u ʊ] by Brazilian college learners of English Language Teaching over the four first semesters of their studies. These six vowels are particularly challenging for Brazilians due to the natural difficulty to perceive and produce sounds of an L2 which are very similar yet not contrasted in the learner's L1, as already stated by Flege's Speech Learning Model [8, 9, 10].

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 [3, 4, 11, 14, 19]. Therefore, it is a very challenging task to perceive and produce L2 contrasting sounds that are very close to a single sound of the L1 [8, 9, 10]. This is the case with English vowels [i ɪ ɛ æ u ʊ], which tend to be assimilated by Brazilian learners into the prototypical categories of Brazilian Portuguese [i ɛ u], respectively [16, 2, 20, 18].

Assuming that the process of L2 acquisition is a complex dynamic system [7, 5, 12, 15], 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, where the system finds temporary stability amidst chaos. These states are temporary due to the dynamic nature of such systems, which may move, or even keep moving, from one attractor state to another. That is why Language Acquisition would be more accurately described as Language Development, due to its dynamic, never-ending change in time as the system moves through different attractor states.

Some attractor states require more energy for the system to move away from, and that is why some learners need more perturbation to have their systems exit an attractor state, but they can all potentially do so. These perturbations might be language lessons, exposure to the L2, interaction with L2 speakers, experiences abroad, etc., but since the relation between perturbation and movement of the system is non-linear, the effects of language lessons, for instance, might not be immediately seen.

In addition, dynamic systems are complex in the sense that the overall behavior of the system is more than the sum of the behavior of its elements, rather it emerges from the iterative interaction of the many elements that make up the system within themselves and with the environment. This makes the L2 learning experience extremely idiosyncratic, for each system (learner) will behave differently at different moments of their developmental route.

This dynamic and idiosyncratic nature of L2 phonological development is what makes dynamic systems better examined in a longitudinal study that, besides looking into group patterns, also analyzes individual routes of development [17, 13, 21, 6], which is the contribution this study attempts to make.

2. METHOD

The data analyzed is the production of the six vowels by 10 Brazilian college students majoring in English Language Teaching. Participants were all male, aged 18-20, native speakers of Brazilian Portuguese, with no experience in an English-speaking country, no experience learning foreign languages other than English, and with no extracurricular English lessons experience (in Brazil English lessons are mandatory in middle and high school, but with the sole focus on reading - teenagers who wish to learn how to speak English either go to a private language institute or study on their own). They had heterogeneous levels of general language proficiency, with some learners at a very basic level, and others with fluent conversational abilities by having studied on their own through podcasts, videos and music. They took five mandatory courses per semester, each one with a total of 64 hours. In the first semester, one course is taught in English and the other four in Portuguese. From the second semester on, all courses are taught entirely in English. In the third semester, they take a mandatory English Segmental Phonology course, in which they gain technical knowledge and practice pronunciation of segments, including the target vowels of this study.

Participants were recorded individually at the end of semesters 1, 2, 3 and 4, 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 all monosyllabic CVCs, with most Cs being voiceless plosives, preventing acoustic bias from neighboring segments. The words were "peak", "Pete" and "teak" for [i]; "pick", "Pitt" and "tick" for [ɪ]; "peck", "pet" and "tech" for [ɛ]; "pack", "pat" and "tack" for [æ]; "boot", "poop" and "toot" for [u]; and "book", "put" and "took" for [ʊ].

F1 and F2 values were used to create vowel space plots for individual speakers in order to compare their development over time. F1 and F2 values were Lobanov-normalized in order to calculate the Euclidean Distances (ED) 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. The ED is a measure of dissimilarity that can be used to measure the distance between two points in a cartesian coordinate system, which is the case of the F1- F2 graph [1]. Finally, the EDs were used to fit a mixed-effects model to the data.

3. RESULTS

The first step was to visually inspect individual vowel spaces, comparing the distributions of speakers' vowels in the four different recordings. When two vowels had half or more of their one-standard-deviation ellipses overlapping, they were considered potential candidates of overlapping vowels with no separate phonetic categories; and when less than half of the ellipses overlapped or when they did not overlap at all, they were considered potential candidates of separate vowel categories. In order to confirm the status of those potential candidates for separate or overlapping vowels, the ED between each contrasting pair of vowels for each speaker was calculated using Lobanov-normalized F1 and F2 values. In a previous study with the same method of data collection and analysis [16], the EDs between the mean formant values of a group of 10 native speakers of American English were .46 for [i ɪ], .38 for [ɛ æ] and .33 for [u ʊ]. Therefore, in this study, the potential separate pairs of vowels were in fact considered separate phonetic categories only if their EDs were of at least 0.3. It is based on these two criteria that table 1 shows in which recordings there is a contrast between the target vowels for each speaker.

As can be seen, there are all types of developmental routes, from a learner that did not develop separate vowel categories at all; to those who developed along the way, especially after taking the English Segmental Phonology course (between recordings 2 and 3); and those who created new phonetic categories but then lost them. From the 10 learners, 7 already had separate vowel spaces for the pair [i ɪ] in recording 1, and the other 3 did not develop these categories in the other three recordings. For the [ɛ æ] pair, only one student already had separate phonetic categories for them in recording 1, three learners developed separate categories for them in recording 3 and kept them in recording 4, and two of them produced them as separate vowels in recording 3 (right after taking the English Segmental Phonology course) but not anymore in recording 4. For the high back vowels, two learners produced them separately in recordings 1 through 3 but not in 4, three learners created separate phonetic categories for them along the way, and only one already had them separate from recording 1 onwards. Only three learners got to recording 4 with separate phonetic categories for all three pairs. The column with most YES's is for the pair [i ɪ] and the one with fewest is the one for [ɛ æ].

Lastly, as an attempt to look at a general vowel development index for each learner and for the group as

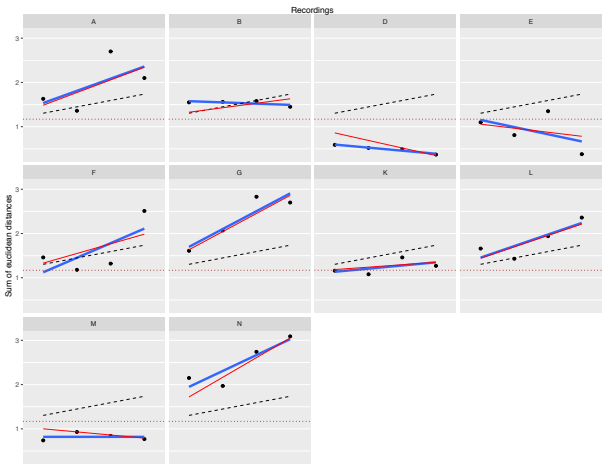
Table 1: Occurrence of separate phonetic categories for target pairs of vowels by speaker and recording

Speaker	Recording	[i ɪ]	[ɛ æ]	[u ʊ]
A	1	**YES**	no	no
	2	**YES**	no	no
	3	**YES**	**YES**	**YES**
	4	**YES**	**YES**	**YES**
B	1	**YES**	no	no
	2	**YES**	no	no
	3	**YES**	no	no
	4	**YES**	no	no
D	1	no	no	no
	2	no	no	no
	3	no	no	no
	4	no	no	no
E	1	no	no	**YES**
	2	no	no	**YES**
	3	no	no	**YES**
	4	no	no	no
F	1	**YES**	no	no
	2	**YES**	no	no
	3	**YES**	**YES**	no
	4	**YES**	**YES**	**YES**
G	1	**YES**	no	no
	2	**YES**	no	**YES**
	3	**YES**	no	**YES**
	4	**YES**	no	**YES**
K	1	**YES**	no	no
	2	**YES**	no	no
	3	**YES**	**YES**	no
	4	**YES**	no	no
L	1	**YES**	no	**YES**
	2	**YES**	no	**YES**
	3	**YES**	**YES**	**YES**
	4	**YES**	no	no
M	1	no	**YES**	no
	2	no	**YES**	no
	3	no	**YES**	no
	4	no	**YES**	no
N	1	**YES**	no	**YES**
	2	**YES**	no	**YES**
	3	**YES**	**YES**	**YES**
	4	**YES**	**YES**	**YES**

a whole, the sum of the EDs of the three target pairs of vowels was used to fit a mixed-effects model to the data. The expectation was that learners would increase their distances as they advanced in time in their studies. In the model, the fixed effects were the intercept and the slope of the trend for the population of all 10 learners (from a linear model of sum of EDs by recording), and the random effects were the deviations in intercept and in slope that each subject's own trend had from the population values (based on [1]).

In Figure 1 the four black dots are the sums of EDs for each recording, and the blue thicker line is the tendency line for each speaker from a simple within-subject linear model. Not all speakers had a positive correlation between the sums of EDs and time. In theory, learners should increase the distances between contrasting vowels as they advance in their study of English, but only six of them ended up with a positive correlation, some of which were not significantly steep.

Figure 1: Mixed-Effects Model of the sum of Euclidean Distances of the 3 pairs of vowels by recording for each speaker.



The black dashed line, which is repeated in every individual plot, is the general tendency of the group, also created from a simple linear model. It favors the hypothesis that learners should increase their distances with time of study, with a positive correlation, but the slope was not significant. The model estimated an initial ED of 1.16 (Intercept) and an increase of 0.144 ED per recording (Slope), but with a p-value for the slope of 0.163 ($r^2 = 0.05$).

The red thinner line is the result of the mixed-effects model fitted for each speaker's data. It shows an increase in ED per recording for 7 speakers, but there was a lot of variance among speakers. The standard deviation (SD) of the intercept in the random effects, which estimates subject-to-subject variation in the intercept, was 0.16. Similarly, the SD of the slope in the random effects was 0.21, and the one for the residuals (which represents the expected scatter around the fitted lines for each subject) was 0.3 (AIC = 57.1; BIC = 65.5).

Lastly, the dotted line, with no slope and repeated in all individual plots at 1.17, marks the sum of the EDs from mean F1-F2 values of a group of 10 native speakers of American English [16]. This serves as a reference, showing that the three learners with negatively-correlated lines for the mixed-effects model (speakers D, E and M) had sums of EDs below that of the group of native speakers.

Most learners produced their vowels with EDs greater than those of the group of native speakers (above the fixed dotted line). This does not mean that they necessarily produced vowels in separate phonetic categories because, in many cases, the one-standard-deviation ellipses in their vowel spaces were still overlapping, which did not happen with the group of native speakers. This means that at some point in their developmental routes, the learners were able to produce some of the target words with distinct vowel categories, but not all of them, or not all of the time, resulting in a lot of variance and thus large ellipses in their vowel spaces, whereas the native speakers were able to maintain their vowel categories completely separate (with ellipses far from each other) at a smaller ED.

4. DISCUSSION

There was a lot of variability in the development of learners. Among the productions of all learners, there was a total of 11 vowel contrasts already present in recording 1, eight new contrasts were created along the four recordings (six of them right after the English Segmental Phonology course), and four were "lost". Under a Dynamic Systems Theory (DST) perspective of language development, this comes as no surprise, as each student is a dynamic system undergoing a process of language development, 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 learners to be at the same stage, even if they take a placement test. Learners arrive at college with different experiences in the L2, levels of motivation, quantity and type of exposure to the L2, just to mention a few individual variables.

The contrasts apparently "unlearned" at some point reveal the non-linear nature of language development, showing that the system is constantly moving towards attractor states. The non-linearity between cause and effect also account for the fact that not all students immediately created new phonetic categories after taking the English Segmental Phonology course. It is possible that later on, and triggered by other perturbations of their systems, those learners that showed no immediate effect will move their systems away from the attractor states of the prototypical L1 vowel categories.

Finally, the results section attempted to categorize students' productions into "contrasting vowels" and "no contrasting vowels'. However, language development is not categorical, but gradient in nature. 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 and followed for the categorization of the results. Nevertheless, under DST, 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 1 produced the contrasts equally overlapped, which was not the case. Some students moved their vowels apart, just not enough to fulfill the pre-established criteria. Likewise, not all speakers with contrasting vowels in Table 1 produced them equally well. Some produced them in the threshold of the criteria, whereas others produced vowels truly separated, with the ellipses far from touching each other. There was variation even within the same speaker. Speakers F, G and K, for instance, all marked with separate categories for [i ɪ], produced contrasts way more separate in the last two recordings.

5. CONCLUSION

The main goal of this paper was to analyze the development of six English vowels by Brazilian college learners of English Language Teaching throughout the four first semesters of their studies. This was achieved by analyzing the creation of new phonetic categories for the target vowels and the developmental route of each learner through visual inspection of vowel spaces, calculation of EDs between contrasting vowels, and the creation of a mixed-effects model on the sum of EDs by recording.

The analyses showed a lot of variability in the development of the target vowels by the learners, which was expected under DST. Many learners developed new phonetic categories throughout these four first semesters, and more phonetic contrasts are expected to develop as they continue their studies. Future investigation of these data will include an analysis of durational patterns as well as analysis of a less monitored production (reading a text). Future studies of this nature could also include perceptual studies as an attempt to witness the emergence of both perceptual and productive vowel categories.

ACKNOWLEDGEMENT

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

6. REFERENCES

[1] Bates, D. M. 2010. *lme4: Mixed-effects modeling with R*. Berlin.

[2] 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. *Interspeech*. Citeseer.

[3] Bybee, J. 2003. *Phonology and language use* volume 94. Cambridge: Cambridge University Press.

[4] Cristófaro Silva, T. 2003. Descartando fonemas: a representação mental da fonologia de uso. In: Hora, D., Collischonn, G., (eds), *Teoria linguística: fonologia e outros temas*. João Pessoa: Editora Universitária.

[5] de Bot, K. 2008. Introduction: Second language development as a dynamic process. *The Modern Language Journal* 92(2), 166–178.

[6] de Bot, K., Larsen-Freeman, D. 2011. Researching second language development from a dynamic systems theory perspective. In: *A dynamic approach to second language development*. John Benjamins 5–24.

[7] de Bot, K., Lowie, W., Verspoor, M. 2007. A dynamic systems theory approach to second language acquisition. *Bilingualism language and cognition* 10(1), 7.

[8] Flege, J. E. 2007. Language contact in bilingualism: Phonetic system interactions. *Laboratory phonology* 9, 353–382.

[9] Flege, J. E., Munro, M. J., MacKay, I. R. A. 1995. Factors affecting strength of perceived foreign accent in a second language. *The Journal of the Acoustical Society of America* 97(5), 3125–3134.

[10] Flege, J. E., Yeni-Komshian, G. H., Liu, S. 1999. Age constraints on second-language acquisition. *Journal of memory and language* 41(1), 78–104.

[11] 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.

[12] Larsen-Freeman, D. 1997. Chaos/complexity science and second language acquisition. *Applied linguistics* 18(2), 141–165.

[13] Larsen-Freeman, D., Cameron, L. 2008. *Complex systems and applied linguistics*. Oxford: Oxford University Press.

[14] Leather, J. 2003. Phonological acquisition in multilingualism. In: Mayo, M. d. P. G., Lecumberri, M. L. G., (eds), *Age and the Acquisition of English as a Foreign Language*. Clevedon: Multilingual Matters chapter 2, 23–58.

[15] Lima, R. M., Jr. 2013. Complexity in Second Language Phonology Acquisition. *Revista Brasileira de Lingüística Aplicada* 13(2).

[16] Lima, R. M., Jr. 2015. A influência da idade na aquisição de seis vogais do inglês por alunos brasileiros. *Organon* 30(58), 17.

[17] Lima, R. M., Jr. 2016. A necessidade de dados individuais e longitudinais para análise do desenvolvimento fonológico de l2 como sistema complexo. *ReVEL* 14(27), 203–225.

[18] Nobre-Oliveira, D. 2007. *The effect of perceptual training on the learning of English vowels by Brazilian Portuguese speakers*. PhD thesis.

[19] Pierrehumbert, J. 1990. Phonological and phonetic representation. *Journal of phonetics* 18(3), 375–394.

[20] Rauber, A. S. 2006. *Perception and production of English vowels by Brazilian EFL speakers*. PhD thesis Florianópolis.

[21] Verspoor, M., de Bot, K., Lowie, W. 2011. *A dynamic approach to second language development: Methods and techniques* volume 29. Amsterdam: John Benjamins Publishing.

[1] The Euclidean Distances were calculated using $\sqrt{(F1.A - F1.B)^2 + (F2.A - F2.B)^2}$, where A and B represent two different vowels.