

Gomes, M. (2016). Brazilian English x Brazilian Portuguese: A dynamic approach for the analysis of diphthongs in forensic contexts. In J. Levis, H. Le, I. Lucic, E. Simpson, & S. Vo (Eds). *Proceedings of the 7th Pronunciation in Second Language Learning and Teaching Conference*, ISSN 2380-9566, Dallas, TX, October 2015 (pp. 108-118). Ames, IA: Iowa State University.

BRAZILIAN ENGLISH x BRAZILIAN PORTUGUESE: A DYNAMIC APPROACH FOR THE ANALYSIS OF DIPHTHONGS IN FORENSIC CONTEXTS

[Maria Lucia de Castro Gomes](#), Universidade Tecnológica Federal do Paraná

This research proposes different objectives and purposes from traditional analyses about special characteristics of the pronunciation of Brazilian speakers of English. Instead of the educational context, the arena is the forensic context, where a Brazilian would be the suspect of a crime and, supposedly, his/her voice in English would be the evidence. The aim of this article is to demonstrate preliminary results of an experiment using acoustic analysis, traditionally used for speaker comparison in forensic phonetics. The voices of Brazilians and Americans were recorded speaking English and Portuguese, and the diphthong [aɪ] was compared in three perspectives – inter-speaker, intra-speaker and inter-language. Preliminary results demonstrate that Brazilians do pronounce diphthongs in a way that makes them different from American native speakers, with values of F2 always higher in the beginning and in the end of the diphthong. As for duration, Brazilians present a greater relative duration of the transition (approximant) when compared to the vowel nucleus.

INTRODUCTION

Forensic phonetics is a branch of applied linguistics that provides knowledge and methods of analysis in the context of criminal investigations involving the use of spoken language. According to Nolan (1997), “since the recording of speech has become commonplace, law enforcers and courts have increasingly called on specialists to give opinions on speech samples, either in court or at the investigation stage” (p.746). The most common activities performed by professionals in this area are speaker comparison for identification, speaker profiling, determination of origin, design of voice line-ups, analysis of background noise and analysis of sound edition. With forensic phonetics as the research arena, this project explores a context in which a Brazilian would be the suspect of a crime and his/her recorded voice would be the evidence.

METHODS

This work presents preliminary results of an experiment on the characteristics of Brazilian speakers of English in the forensic context, using speaker comparison as a methodology. The segment for acoustic analysis was the diphthong. The participants were three Brazilians and three American female speakers, who are bilingual in Brazilian Portuguese and American English. Before the recording session, they received instructions on the research and signed to permit the use of their voice in the analyses.

The participants' voices were recorded reading carrier sentences¹ that contained the diphthong [ai] in target words, both in English and Portuguese. The words in English were *pie*, *buy*, *lie* and *high*, and the words (or nonce words) in Portuguese were *pai*, *bai*, *LAI* and *Rai*². The carrier sentences were "Say _____ to me" and "Diz _____ também".

The three Brazilian participants were university students at the time of the study, two 19 year-olds and a twenty year-old. All three of them are from Curitiba, the capital of the southern state of Paraná, and had always lived there. They have Portuguese as their native language and are fluent in English (they said they spoke English every day, mainly at the university). One of them is also fluent in German. The three American participants – two 23 year-olds and a 25 year-old – were English Teaching Assistants sponsored by the Fulbright Program at a university in Curitiba. They were born in the USA, New York (NY), Los Angeles (CA), Lyndhurst (OH), and the three had lived in other cities of the United States. One of them, of Japanese ancestry, had also lived in Tokyo and Honolulu. Another participant of eastern ancestry, Vietnamese, had never lived outside the US. They all had English as their native language, were fluent in Portuguese, and in other languages as well. One was fluent in Japanese and Spanish, the other in Vietnamese, and the third, in French.

The hypothesis that motivated this research is that Brazilian speakers of English may have special features that might indicate their origin in speaker comparison cases in forensic contexts. The choice for the diphthong as the object of analysis came from the potential of vowels in forensic analysis, as advocated by several authors (Rose, 2006; Mcdougall, 2002, 2004). The measurement of F2 was inspired by Nolan and Grigoras (2005), who claimed that the dynamics of the trajectory of formants in diphthongs is very useful in speaker recognition. According to the authors, the second formant presents a clearer trajectory. Haupt and Seara (2012) found a clear transition region between the two vowel targets of the diphthong [ai], observed in the F2 trajectory. Grigoras and Nolan (2005) also state that F2 frequency analyses may be more interesting than the F1 frequency and the F1 values may present errors in telephone recordings due to bandwidth problems.

Preliminary results

The results of the F2 measure come from five repetitions of four words in carrier sentences, in Portuguese and English, pronounced by each of the six participants. For duration, only five repetitions of the words *pai* and *pie*, by the six participants, will be described. Table 1 demonstrates F2 measures of the start and end points of the diphthongs.

¹ There were two other tasks (text reading and a short conversation), but their results won't be shown in this article.

² From those four words, only *pai*, which means *father*, is a real word in Portuguese. The others were invented and inserted in texts. The texts were designed to create a context, so we could have the same pronunciation in English and Portuguese. Brazilians usually use the term *bai*, *bai* to say goodbye, copying the English expression *bye-bye*, so *bai* was inserted in a text, under this meaning. LAI is an acronym, meaning *Lei de Acesso à Informação* – Law on the Access to Information, and the text had the acronym with that meaning. Finally, *Rai* was in the text as a short name for a girl named *Raiana*.

Table 1

F2 measures at the start and end points of the diphthong

Word	Participant	F2-START	F2-END	Word	Participant	F2-START	F2-END
	BRAZ	1490.667	2383.867		BRAZ	1560.867	2521.267
PAI	AM	1432.333	2275.467	PIE	AM	1292.333	2362.267
	BRAZ	1541.6	2358.933		BRAZ	1557.4	2585.467
BAI	AM	1431.733	2314.933	BUY	AM	1340.6	2346.733
	BRAZ	1411.467	2305.533		BRAZ	1364.2	2556.33
LAI	AM	1359.4	2286.667	LIE	AM	1272.67	2302.27
	BRAZ	1651.067	2297.2		BRAZ	1641.067	2565.133
RAI	AM	1559.4	2326.933	HIGH	AM	1399.867	2365.8

Statistical analyses of the data are yet to be done, but it is already possible to see some tendencies in the results. As seen in Table 1, the means of F2 values, either at the beginning or at the end of the diphthong (except for *Rai* F2-Start in Portuguese), had higher values in the productions by the Brazilian subjects, both in Portuguese and English words. Figure 1 provides a better view of the results.

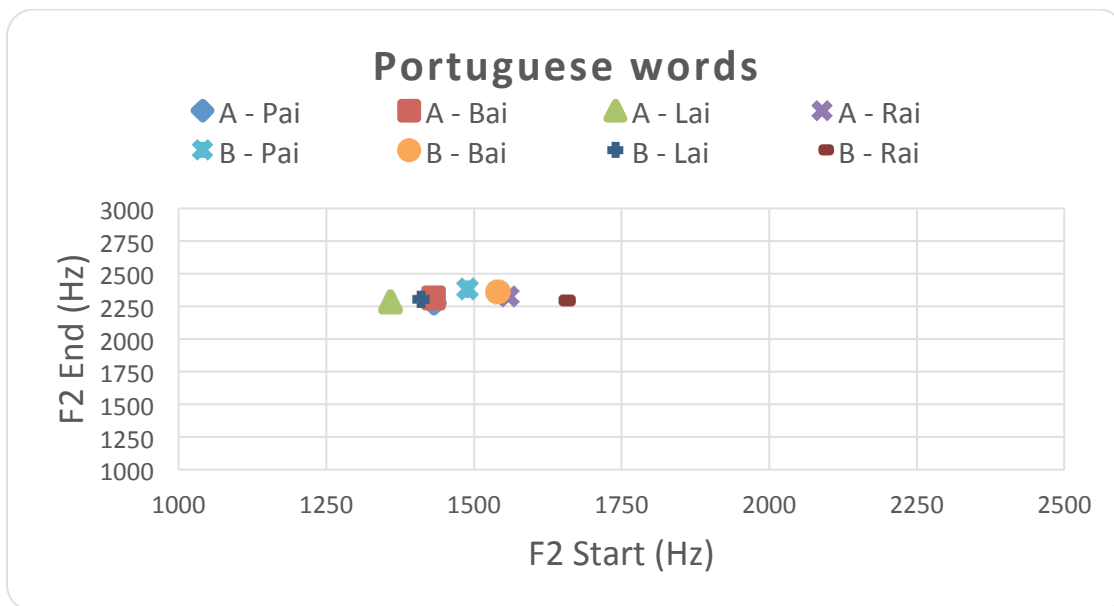


Figure 1. F2 converging point at the beginning and end of the Portuguese diphthongs [ai] produced by the Americans (A) and the Brazilians (B) – Portuguese words.

For the words in Portuguese, we can see that the mean F2 at the beginning of the diphthong is between 1,359 Hz (in *LAI* by the Americans) and 1,651 Hz (in *Rai* by the Brazilians). On the other hand, F2 at the end of the diphthong is between 2,275 Hz (in *pai* by the Americans) and 2,383 (in *pai*, by the Brazilian). We can see a larger difference in F2 values from the start point to the end point of the diphthong. We can also see that there may be a phonological influence, not analyzed here. However, we can see that at the start point, F2 values are higher when the preceding consonant is a plosive. And we can also see that F2 values in the beginning are higher when the diphthong is preceded by [h] than when preceded by [l], when pronounced by both the Brazilians and the Americans. Comparing the start point of F2 produced by both groups, word by word, one can clearly see the converging point always further to the right when produced by the Brazilians. The difference between Brazilian and American productions is also present in the English words.

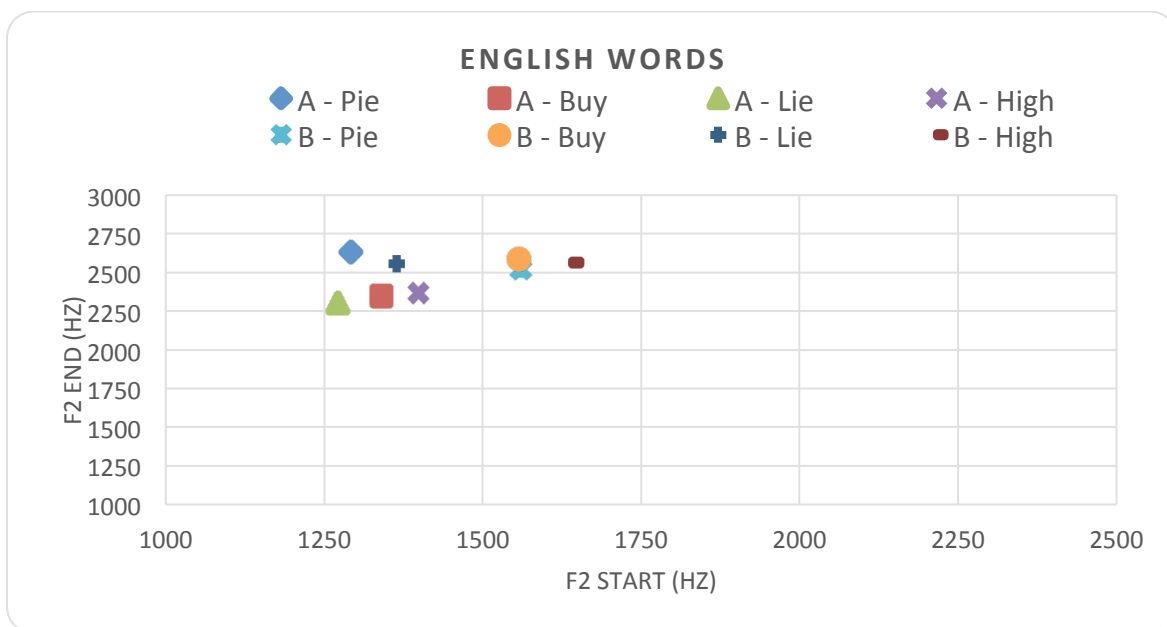


Figure 2. F2 converging point at the beginning and end of the Portuguese diphthongs [ai] produced by the Americans (A) and the Brazilians (B) – English words.

In Figure 2 we can also observe that the converging points of F2 produced by the Brazilians have higher F2 values. The mean F2 value at the start point of the diphthong is between 1,272 Hz (in *lie* by the Americans) and 1,641 Hz (in *high* by the Brazilians). The mean F2 value at the end of the diphthong is between 2,302 Hz (in *lie* by the Americans) and 2,585 (in *buy* by the Brazilians). Comparing word-by-word, it is possible to observe that the converging points of the two F2 measures are always at a higher position at the Brazilians' graph data, the values of the end point of F2 being about 200 Hz higher. According to Yavas (2011, cited by Woolums, 2012), in order to qualify as significant, the difference in F2 must be 200 Hz.

As for the Brazilians' productions of Portuguese and English words, there is a tendency for higher value of F2 at the end of the diphthongs in the words in English, as can be seen in Figure 3. Except for the word *pie*, which had a difference of 137.4 Hz higher in the F2 value at the end of the diphthong, all the other differences were higher than 200 Hz.

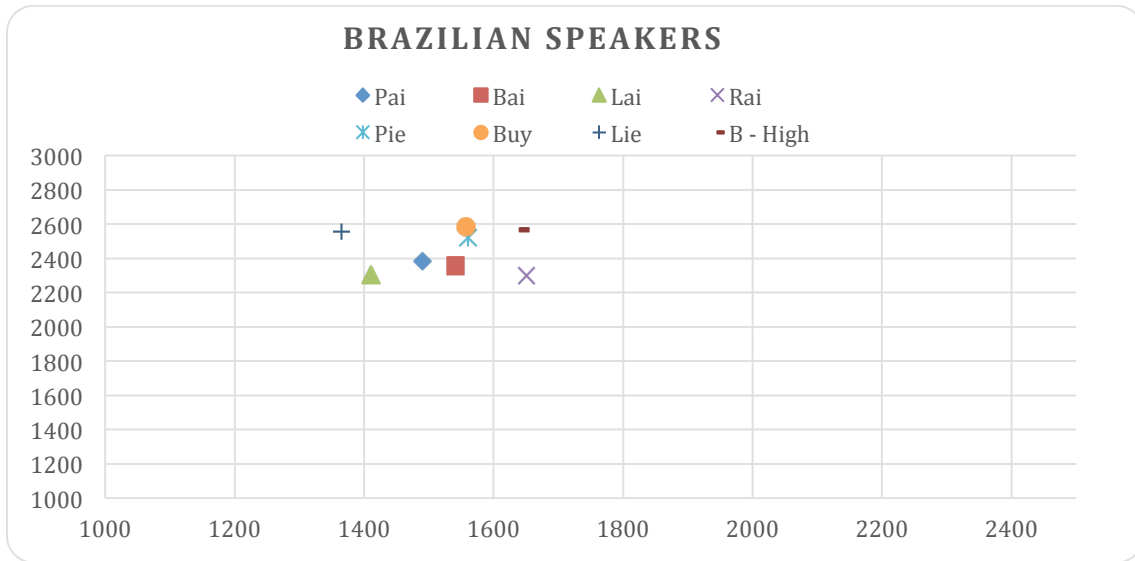


Figure 3. F2 converging point at the beginning and end of the diphthong [aɪ] produced by the Brazilians in the words in Portuguese and in English.

Figure 3 demonstrates that the converging point of F2 at the beginning and at the end of the diphthong is always higher for the words in English than the words in Portuguese. After comparing the diphthongs produced by the Brazilian and American participants in the words in English and Portuguese, and between the diphthongs in the words in Portuguese and in English produced by the Brazilians, the pair of words *pai/pie* was chosen to establish a comparison between the three informants of each nationality. Checking some dispersion may indicate important linguistic variations, mainly due to dialect differences. The first charts compare the words pronounced in the native language, i.e. Portuguese by the Brazilians and English by the Americans.

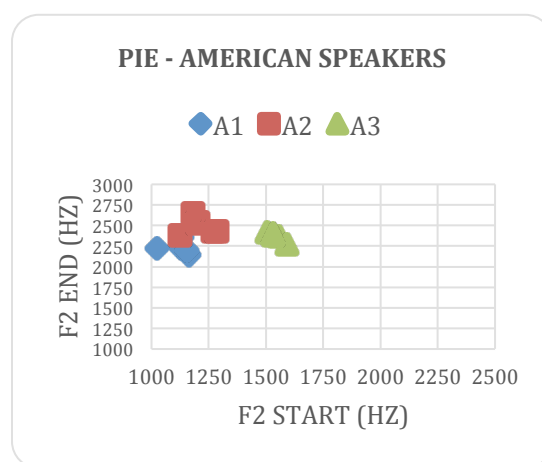


Figure 4 and 5. F2 converging point at the beginning and end of the diphthongs [aɪ] by the Brazilians (Figure 4) in the word *pai* and by the Americans (Figure 5) in the word *pie*.

The charts indicate that the productions of the Brazilians exhibit less variation than those of the Americans. There may be less influence of dialectal variation since the Brazilian subjects were all born in the same city, while the Americans were all born in different regions of the US. Another fact that may be relevant is that the diphthong [ai] is undergoing raising in some Midwestern portions of the US (Woolums, 2012).

Analyses of the productions of the same pair of words as a foreign language, by both groups of subjects, is displayed in Figures 6 and 7.

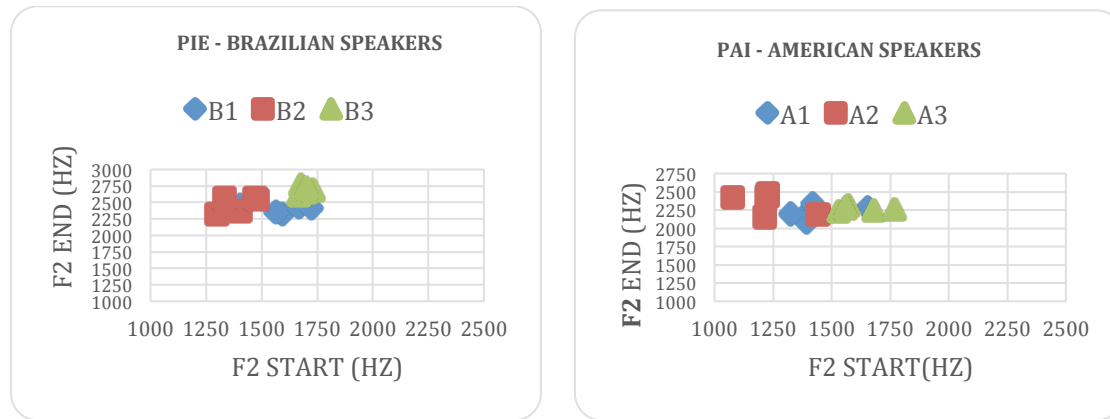


Figure 6 and 7. F2 converged points at the beginning and end of the diphthongs [ai] by Brazilians respondents (Figure 6) in the word *pie*, and Americans (Figure 7) in the word *pie*.

Duration

Next, we analyze the duration of the pair of words *pie/pai* in detail. For the duration of the vowel, the start and end points were the first and last periods that had considerable amplitude. From the point where the trajectory of formants (F1 and F2) changed, we considered the final point of the vowel and the start of the approximant. The end point of the approximant was the last pulse of the waveform. Figure 8 demonstrates an example of duration measurement using PRAAT. The first tier shows the waveform, the second tier shows the spectrogram, in the third tier we can see the word, and finally in the fourth, we see the labels of the segments. Although the consonant is not a focus of analysis here, it was also labeled, as it may be important in the word duration, especially if we consider the aspiration in the English word.

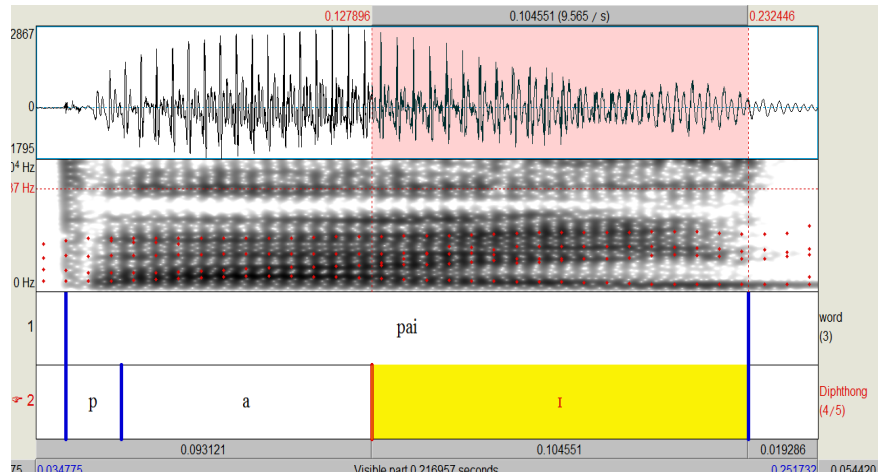


Figure 8. Word *pai* produced by Subject B3.

For the calculation of the relative duration of the vowel and the approximant, we divided the absolute duration of each by the absolute duration of the word, and the result was multiplied by 100, following Silva (2014). We can see, in the tables below, the mean values of the absolute duration of the word (in seconds) and the relative duration of the vowel and the approximant of the words *pai* and *pie* produced by the Brazilians and the Americans.

Table 2

Absolute duration of the word and relative duration of the vowel and the approximant by the Brazilians and the American, and standard deviation in the Portuguese word pai.

INF	Word dur.	Vowel rel.dur	Approx rel. dur.	INF	Word dur.	Vowel rel.dur	Approx rel. dur.
B1				A1			
	0.18	29.42	57.13	A2	0.28	42.34	45.95
B2				A3	0.24	40.30	45.61
	0.19	32.04	59.18	MEAN	0.25	40.37	37.24
B3				SD	0.26	41.00	42.93
	0.24	37.33	58.86		0.015	0.94	4.02
MEAN	0.20	32.93	58.39				
SD	0.02	3.29	0.90				

Table 3

Absolute duration of the word and relative duration of the vowel and the approximant by the Brazilians and the American, and standard deviation in the English word pie.

INF	Word dur.	Vowel rel.dur	Approx rel. dur.	INF	Word dur.	Vowel rel.dur	Approx rel. dur.
B1				A1			
B2	0,26	35,42	43,17	A2	0,30	31,84	36,24
B3	0,24	42,49	51,20	A3	0,40	34,26	39,58
MEAN	0,29	34,58	51,31	MEAN	0,34	36,78	39,18
	0,26	37,50	48,55		0,35	34,30	38,33
SD	0,011	3,53	4,01	SD	0,041	2,016	1,489

Table 2 shows the mean duration of the word *pai* [pai] in Portuguese, and Table 3 shows the values of the word *pie* [pai] in English. On the left side of the two tables are the values of the productions of the Brazilian subjects, and on the right, the ones of the American subjects. Comparing the absolute duration of the word, there is a higher mean duration of the word in English than in Portuguese by both groups of subjects. On the other hand, the mean duration of the word is higher when produced by the Americans, in both languages, Portuguese and English. The difference between the relative duration of the vowel and the approximant is higher when produced by the Brazilians, mainly in Portuguese.

DISCUSSION

Though very preliminary, the above data indicate that the hypothesis that Brazilians have special characteristics when they speak English is likely to be confirmed. F2 measures of the diphthong [ai] in the Brazilian subjects' productions reveal a different characteristic, both in comparison to the production of the Americans in English, and when compared to their own productions of words in English and in Portuguese.

It is important to acknowledge several factors that could certainly be at stake in the data analysis, mainly the consonantal context and dialectal variation. Concerning consonantal context, the influence of the consonant that precedes the diphthong is evident in the F2 values, especially at the beginning of the diphthong. Figure 3 clearly shows that influence, but it also points to a tendency for vowel fronting by the Brazilian subjects, mainly in the English words, since higher F2 values indicate greater fronting.

A second factor that might influence the results and should be taken into account is the dialectal variation of diphthongs in English. Jacewicz and Fox (2013) analyzed [ai] in three different dialects of North American English and claim that the duration of diphthongs varies significantly in different regions of the United States and these temporal factors may have profound effects on the trajectory of the formants along the vowel. Jacewicz and Fox (2013) claim that there is a

complex interaction between the duration of the vowel, the consonant context and the speaking style, influencing the frequency of formants.

Jacewicz and Fox (2013) was dedicated to temporal issues of vowels, and reveals that the duration differences due to consonant context interact with dialectal variation. One of the regional dialects analyzed was the one produced in Ohio, birth region of one of the participants in this study (A3). She had very different results from the other two American subjects. Her F2 values, especially on the starting point of the diphthong, were closer to the F2 values of the Brazilian subjects. The difference between the F2 values of the Brazilians and the Americans could have been larger if that informant was not in the group. Figure 4 shows that the converging points of A3, in the productions of the word *pie*, is closer to the Brazilians' data. As mentioned above, work by Woolums (2012) also demonstrates [aɪ] raising in some US dialects.

Analysis of the duration in these preliminary results may also indicate a difference in the pattern of production between the Brazilians and Americans. Tasko and Greilick (2010) analyzed the duration of American diphthongs, comparing clear and conversational speech, and concluded that although the duration of diphthongs is greater when spoken clearly, the ratio is approximately the same as in conversation, i.e. 46% in articulated speech and 45% in the conversation. Analyzing the production of diphthongs at different speeds, they tried to explain the phenomena outlined in the gestural overlap model designed by Tjaden and Weismer (1998 *apud* Tasko & Greilick, 2010), saying that “shorter overall durations can be

achieved by increasing the degree of gestural overlap, whereas increasing duration would be achieved by reducing the degree of gestural overlap” (Tasko & Greilick, 2010, p.97).

In this research, the Brazilian subjects' data present a greater relative duration of the approximant when compared to the vowel, demonstrating a greater overlap of the approximant. The Americans distributed the vocalic segments in a more balanced way.

ABOUT THE AUTHOR

Maria Lúcia de Castro Gomes is a PhD in Linguistics and works as a professor at UTFPR – Universidade Tecnológica Federal do Paraná, in Brazil. She coordinates the course of Language and Literature at that institution. Her research focus is on phonetics and phonology of English and Portuguese, and forensic phonetics.

REFERENCES

- Becker, M. R. (2012). O construto “Inteligibilidade” da língua Inglesa sob o paradigma de língua franca. In *Anais do X Encontro do CELSUL – Círculo de Estudos Linguísticos do Sul* (pp. 1-10). Cascavel: Unioeste.
- Becker, M. R. (2014). Globalização, Inglês como Língua Franca e Inteligibilidade. In: In Brawerman-Albini, A & Gomes, M. L. C. (Eds) *O jeitinho brasileiro de falar inglês –*

- pesquisas sobre a pronúncia do inglês por falantes brasileiros* (pp. 223-238). Campinas: Pontes Editora.
- Brawerman-Albini, A., & Gomes, M. L. C. (2014a) Aquisição de inglês como segunda língua perspectivas teóricas para o ensino de pronúncia. In A. Brawerman-Albini & M. L. C. Gomes. (Eds). *O jeitinho brasileiro de falar inglês – pesquisas sobre a pronúncia do inglês por falantes brasileiros* (pp. 19-32). Campinas: Pontes Editora.
- Brawerman-Albini, A., & Gomes, M. L. C. (2014b) *O jeitinho brasileiro de falar inglês pesquisas sobre a pronúncia do inglês por falantes brasileiros*. Campinas: Pontes Editora.
- Cruz, N. F. C., & Pereira, M. A. (2006). Pronúncia de aprendizes brasileiros de inglês e inteligibilidade: um estudo com dois grupos de ouvintes. *Revista Virtual de Estudos da Linguagem – ReVEL*, 4(7) 1–26.
- Cruz, N. F. C. (2014). Inteligibilidade e o ensino de pronúncia do inglês para brasileiros. In A. Brawerman-Albini & M. L. C. Gomes. (Eds). *O jeitinho brasileiro de falar inglês – pesquisas sobre a pronúncia do inglês por falantes brasileiros* (pp. 239-252). Campinas: Pontes Editora.
- Crystal, D. *English as a global language*, 2nd ed. (2010). Cambridge: Cambridge University Press.
- Gold, E., & French, P. (2014). Commonly used forensic speaker comparison parameters: expectations, misconceptions and empirical findings. *International Association for Forensic Phonetics and Acoustics 2013 Annual Conference*. Retrieved from <http://www.iafpa.net/conf.htm>.
- Gomes, M. L. C., Brawerman-Albini, A., & Engelbert, A. P. (2013). *The perception of vowel epenthesis and stress in an English as a Lingua Franca context. Proceedings of the International Symposium on the Acquisition of Language Speech*, Concordia Working Papers in Applied Linguistics.
- Graddol, D. (2006). *English Next: Why global English may mean the end of ‘English as a Foreign Language’*. Retrieved from www.britishcouncil.org/learning-research
- Haupt, C., & Seara, I. C. (2012). Caracterização acústica do fenômeno de monotongação dos ditongos [ai, ei e oi] no falar florianopolitano. *Linguagem & Ensino*, 15(1), 263-290.
- Jacewicz, E., & Fox, R. A. (2013). Cross-dialectal differences in dynamic formant patterns in American English vowels. In *Vowel Inherent Spectral Change*, ed. by Geoffrey S. Morrison and Peter Assmann, 177-198. New York: Springer-Verlag.
- Mcdougal, K. (2002). Speaker-specific formant dynamics: An experiment on Australian English /aɪ/. *Speech, Language and the Law*, 11(1).
- Mcdougal, K. (2006). Dynamic features of speech and the characterization of speakers: towards a new approach using formant frequencies. *Speech, Language and the Law* 13(1).
- McCrocklin, S. (2012). *The role of word stress in English as a lingua franca. Proceedings of the 3rd Annual Pronunciation in Second Language Learning and Teaching* (pp. 249-254). TESL/Applied Linguistics Iowa State University.

- Munro, J. M., Derwing, T. M., & Holtby, A. K. (2012). *Evaluating individual variability in foreign accent comprehension. Proceedings of the 3rd Annual Pronunciation in Second Language Learning and Teaching Conference* (pp. 233-239). TESL/Applied Linguistics Iowa State University.
- Nolan, F. (1995). Can the definition of each speaker be expected to come from the laboratory in the next decades? *ICPhS 95 Stockholm*, 3, 130-137.
- Nolan, F. (1997). Speaker recognition and forensic phonetics. *The Handbook of Phonetic Sciences* (pp. 744-767). Oxford: Blackwell.
- Nolan, F. & Grigoras, C. (2005). A case for formant analysis in forensic speaker identification. *Speech, Language and the Law*. 12(2), 143-173.
- Rauber, A.S., Watkins, M., Silveira, R., & Koerich, R., Eds. (2010). *The acquisition of second language speech: Studies in honor of Professor Barbara O. Baptista*. Florianópolis: Insular.
- Roberson, A. (2012). *Analyzing comprehensibility among non-native speakers of English: The effect of listener first language background. Proceedings of the 3rd Annual Pronunciation in Second Language Learning and Teaching* (pp. 240-248). TESL/Applied Linguistics Iowa State University.
- Rose, P. (2006). *The intrinsic forensic discriminatory power of diphthongs*. Proceedings of the 11th Australian International Conference on Speech Science & Technology, University of Auckland, New Zealand. December 6-8.
- Schadech, T. S. & Silveira, R. (2014). A inteligibilidade e o ensino de pronúncia: interface entre pesquisa e pedagogia. In A. Brawerman-Albini & M. L. C. Gomes. (Eds). *O jeitinho brasileiro de falar inglês – pesquisas sobre a pronúncia do inglês por falantes brasileiros*. Campinas: Pontes Editora. 253-271.
- Silva A. H. P. Organização temporal de encontros vocálicos no Português Brasileiro e a relação entre Fonética e Fonologia. *Letras de Hoje*, Porto Alegre, v. 49, n. 1, p. 11-18.
- Tasko S. M., & Greilick K. (2010). Acoustic and articulatory features of diphthong production: A speech clarity study. *Journal of Speech, Language, and Hearing Research*, 53, 84–99.
- Woolums, N. (2012) Phonetic manifestations of /aɪ/ raising. *Linguistic Portfolios*, 1, Art.19.