ACOUSTIC AND PERCEPTUAL CHARACTERISTICS OF REDUCED VOWELS PRODUCED BY SPEAKERS OF ENGLISH AS A FOREIGN LANGUAGE

Carina Silva Fragozo

Pontifical Catholic University of Rio Grande do Sul, Brazil
cfragozo@gmail.com

ABSTRACT

This article analyzes the process of vowel reduction in function words by Brazilian speakers of English as a Foreign Language. It compares information obtained by perceptual verification with those obtained by acoustic verification. The sample is composed by sixteen female speakers of English as a FL and one native speaker of English. Data collection was accomplished through an instrument composed by sixty affirmative sentences containing the function words *at, for, from, of* and *to*. The data were analyzed both perceptually and acoustically.

It was found that speech rate, measured by the duration of each sentence, vowel duration and speech register were relevant linguistic factors for the perception and production of reduced vowels.

Keywords: vowel reduction, function words, speech acoustics, speech perception

1. INTRODUCTION

Vowel reduction, which is defined in this paper as the replacement of a full vowel by a schwa [ə], is very common in English and occurs mainly in function words. Native speakers of English usually produce a schwa in function words during natural speech. The use of a full vowel in these words is restricted to cases in which there is an intention of emphasis or when the function word is quoted. However, Brazilian speakers of English as a Foreign Language (FL) tend to produce full vowels in function words even when there is no intention of emphasis, which alters the rhythm of the language and characterizes a foreign accent. Based on this fact, this research aims at analyzing the process of vowel reduction in function words produced by speakers of English as a FL.

According to Crosswhite [2], it is a mistake to attribute a universal pattern for vowel reduction, as each language presents different characteristics of the phenomenon. For Chomsky and Halle [1], the only vowel that is totally reduced is the schwa, but this kind of reduction is not common in languages like Portuguese, for instance. Therefore, both Portuguese and English present vowel reduction, but this phenomenon occurs differently in each language.

In Portuguese, the vowels of function words are usually reduced, but this reduction does not occur from a full vowel to a schwa, as they tend to be reduced to zero (ex.: *d(e)uma* → [ʻduma]).

In English, vowel reduction is essential for the rhythmic organization of the language, and it is common in function words, which usually present a schwa in natural speech. According to Roach [11], it is possible to use full vowels in function words, but this production is more common in foreign speech. Thus, vowel reduction occurs categorically in function words in English, but this process depends on aspects like style and speech rate in Portuguese.

Thus, contrarily to native speakers of the language, Brazilian speakers of English as a FL may produce either full vowels or schwa in function words, as in examples (a) and (b):

(a) Give this book **<u>to</u> John**.
[tʊ] ~ [tu] ~ [tʌ] ~ [tə]
(b) I cook **<u>at</u> home**.
[æt] ~ [ɛt] ~ [at] ~ [ət]

The aim of this paper, then, is to analyze the variation in the production of reduced vowels in function words by speakers of English as a FL and to compare the perception and the production of these vowels, which is accomplished through acoustic and perceptual verifications.

2. METHOD

The sample consisted of seventeen female subjects: sixteen Brazilian speakers of English as a FL, all of them with a minimum of 3 years studying the language, and one native speaker of English, who was used as a small reference sample.

In spite of the fact that vowel reduction is a phenomenon of natural speech, data collection was accomplished through an instrument for relevant reasons. Firstly, because it makes it easier to calculate the occurrences, as all the subjects

produce the same sentences. In addition, because the instrument allows absolute control on the preceding and following contexts, which minimizes the effects of coarticulation in the spectrogram examination. The instrument was composed by twelve affirmative sentences containing the prepositions *at, for, from, of* and *to*, the total of 1020 occurrences to be analyzed acoustically and perceptually. In order to avoid the emphasis on the function words, a part of the sentence that should be produced emphatically was underlined in each sentence, which induced the subjects to produce the prepositions in a weak way. All the sentences in the instrument presented a maximum of two content words after the preposition, so that all the function words were in the same prosodic and syntactic conditions. Thus, the instrument contained sixty sentences presented as the following examples:

a) I'm not good at multiplication
b) Honesty is essential for politicians
c) I bought a very rare stamp from Canada.
d) We bought a tube of toothpaste.
e) Next year I'll travel to California

In order to reduce the effects of writing in the production of the subjects, they were told to read each sentence twice, and then they repeated it by heart, without reading. Only this third production was considered for the analysis, except for a few cases in which the subjects stuttered and the first or the second productions were used.

The digital recorder Olympus VN-120, with a sampling frequency of 16.0 kHz, was used for data collection. The formants of the vowels were measured through the software Praat, version 5.0.47.

After the recordings, the first procedure was to make the perceptual analysis of the data. Each sentence was heard at least three times, so that the vowels could be classified as "full" or "reduced". Then, the vowels were acoustically analyzed through Praat. In order to reduce the effects of coarticulation, only the central portion of the vowel was selected. The selection ranged between 10 and 30 ms because, according to Flanagan [4], the configuration of the vocal tract is in a relatively steady position in this time interval.

Due to the weak position of these vowels, some of them were deleted or produced very fast. Then, we established a minimum of three glottal pulses to consider the values of F1 and F2 of the vowels. If they presented less than three pulses, we classified them a deletion.

For establishing the limits between a full vowel and schwa, we obtained the formant values for female voice of all the vowels that could probably be produced in the prepositions in analysis. We did not take the peripheral vowels from the speakers taking part in the experiment because we aimed at comparing native and non-native productions, so it was necessary to use information from the literature. The values of the English vowels [ɛ æ ʌ ɑ u ʊ], extracted from Yavas [13], and the values of the Portuguese vowels [a ɔ], extracted from Escudero, et. al [3] are presented in Table 1:

Table 1: Formant values of vowels that could be produced in the prepositions *at, for, from, of* and *to*.

POSSIBLE PRODUCTIONS	F1	F2
ɛ	600	2350
æ	860	2050
ʌ	760	1400
a	910	1627
ɔ	442	893
ɔ	590	900
ɑ	850	1200
u	370	950
ʊ	470	1150

The formant values were transferred to tables in the software MATLAB® (MATrix LABoratory, version 6.0.0.88), which calculated the Euclidian distance between the vowels and measured the ones that were closer to schwa and the ones that were closer to the values of the full vowels presented in Table 1. For the statistical treatment of the data, we used the software GoldVarb-X, developed by David Rand and David Sankoff.

3. RESULTS AND DISCUSSION

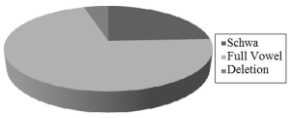
The application rate of schwa in the function words in analysis can be visualized in the following graphs. Graph 1, which corresponds to the results obtained through acoustic verification, shows that schwa occurred in 24,4% of the data (234/960), full vowels occurred in 71,1% of the data (683/960) and deletion occurred in 4,5% of the data (43/960).

As to the perceptual verification, the results indicate that the vowel schwa occurred in 37% of the data (355/960) and full vowels occurred in 63% of the data (605/960), as shown in Graph 2. Occurrences of deletion were not considered perceptually.

We can observe in Graphs 1 and 2 that there are differences in the application rate of schwa in the acoustic verification and in the perceptual verification. According to Johnson [6], the scales of frequency and height of the unit of analysis (such as a computer, for example) are not the same as the auditory system. This partially explains the

fact that the acoustic verification of speech data does not agree with the perceptual verification, as in this study. Besides, the relation between speech perception and production is a relevant aspect to be considered. According to Reets and Jongman [10], the most important acoustic characteristic for the perception of the quality of a vowel is the value of its formant frequencies. However, the authors say that identifying the first two formants of a vowel is not enough for the listener to classify it in an internal category. According to them, the phonetic context, the speech rate and the size of the speaker's vocal tract may influence the frequency pattern of the formants.

Graph 1: Acoustic verification.



Graph 2: Perceptual verification.

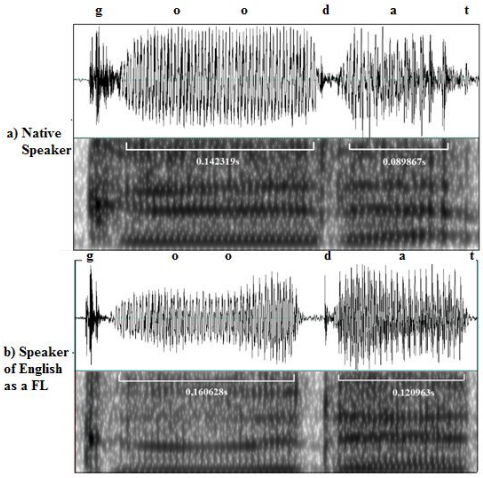
As to the perception of vowels as full or reduced, in addition to the tongue position, acoustically identified by the values of F1 and F2, aspects like *intensity* and *duration*, which is directly related to speech rate, may be determinant for vowel classification in the listener's perspective.

The acoustic intensity, which corresponds to the amount of energy in the vibratory movement, may influence the perception of vowels as full or reduced. In other words, it is possible that when the vowels of function words are produced with a very low intensity they may be heard as reduced vowels, even though its quality is of a full vowel.

For Lieberman and Blumstein [7], duration is one of the aspects that influence the identification of the quality of English vowels. The authors say that the duration patterns for each vowel are relative, because they vary according to speech rate, to the degree of stress of the vowel, and to the context, and they affirm that the listener usually identifies vowels by their relative duration. With regard to the influence of duration on the formant values of vowels, Rosner and Pickering [12] affirm that speech rate does not cause significant alterations in the values of F1 and F2. Due to the fact that schwa is the vowel with the shortest intrinsic duration, as Lindblom [8] says, we believe that when the vowels of function words are produced with a shorter duration, they are heard as a reduced vowel, even if the acoustic verification shows that the *quality* of these vowels, indicated by their formant values, does not correspond to the reduced vowel schwa.

The differences of duration between a full vowel and a reduced vowel can be visualized in Figure 1, in which both oscillograms (and their corresponding spectrograms) represent the expression *good at*, which was present in one of the sentences of the instrument. In (a), we have the production of the native speaker, and in (b), the speech of a highly proficient speaker of English as FL. The duration of the vowels is identified in Figure 1 with the numbers indicated between the oscillogram and the spectrogram, in seconds.

Figure 1: Oscillograms and spectrograms of the expression *good at* produced by a native speaker and by a speaker of English as a FL.



It is possible to observe in Figure 1 that both vowels in (a) present shorter durations than the ones in (b), and that the reduced vowel in *at* is shorter than the full vowel in *good* in both productions. This difference is better illustrated in Table 2:

Table 2: Vowel duration in a full vowel and in a reduced vowel: native and non-native speech.

	Vowel duration in *good*	**Vowel duration in *at***	**Full vowel vs. reduced vowel**
Native	0,142319	0,089867	**58,37%**
Non-native	0,160628	0,120963	**32,79%**
Native vs. Non-native	**12,86%**	**34,60%**	

Table 2 shows that the duration of the reduced vowel in *at* is 58,37% shorter than the full vowel in *good* in native speech, and that this difference is of 32,79% in non-native speech. As to the comparison between native and non-native speech, the table shows that vowel duration in *good* is 12,86% shorter in the native production and the

duration in *at* is 34,60% shorter. This shows that, in addition to the F1 and the F2 of the vowels, duration is also a pertinent element for the analysis of vowel reduction.

By using the software Matlab®, which calculated the Euclidian distance of the vowels, it was possible to identify the productions that were closer to the value of schwa for female voice, that is, F1:550 and F2:1650, as presented in Marusso [9]. In addition, this software identified the quality of all the vowels that were not classified as a schwa, by calculating the formant values of each full vowel that could be produced instead of schwa.

These calculations made it possible to relate the classification of the vowels in the acoustic analysis according to the Euclidian distance and the classification of the vowels between full and reduced in the perceptual analysis. The results show that the phonetic production that most favored the perception of the vowel as reduced was deletion, with a relative weight of 0,951 (39/43). The reduced vowel [ə], with relative weight of 0,798 (155/234), also favored the perception reduced vowels. The vowels [ʊ], with relative weight of 0,697, and [ɛ], with 0,573, also favored the perception of reduced vowels. The vowel [ʌ], with relative weight of 0,501, was not statistically significant, although it is has a very similar tongue position than that of schwa. The vowels [ɔ], [a], [ɑ] did not favor the perception of the vowels as reduced, probably because they are tense and present a longer duration. Finally, the vowel [æ] was heard as a reduced vowel in only 7% of the data (8/115). This may be explained by the fact that this vowel has a longer duration, which facilitates its perception as a full vowel.

These results can be related to the sonority scale of vowels proposed by Hammond [5], represented as follows:

æ>>ɑ>>e.....>>ʌ>>ɪ>>ʊ>>ə

According to the author, low vowels are more sonorous than high vowels, and front vowels are more sonorous than back vowels. In his sonority scale, schwa is the least sonorous and the shortest vowel. The vowels [ɪ] and [ʊ] are also very low in the scale, which agrees with the results in this research, in which the vowel [ʊ] was the one that most favored the perception of reduced vowels, after deletion and schwa. According to the scale, the most sonorous vowels are [æ] and [ɑ], both vowels that did not favor the perception of vowel reduction in this research. Therefore, the relation between the sonority scale and the statistical results in this paper shows that the least sonorous vowels, which are consequently shorter, favor the perception of the vowel as reduced.

Another aspect that confirms the role of speech rate in the perception of reduced vowels is the fact that the native speaker was the subject with the highest percentage of applications of schwa in the perceptual verification (88,3%), but not in the acoustic verification (48,3%). However, among all the subjects, the native speaker was the one who presented the shortest sentence durations, which corroborates the importance of speech rate and, consequently, vowel duration, for the recognition of reduced vowels in the perceptual verification.

4. CONCLUSION

This article showed that there were less applications of schwa in the acoustic verification than in the perceptual one. We found that besides the values of F1 and F2 of the vowels, indicated by the acoustic analysis, the duration and probably the intensity in the production of vowels are aspects that contribute to their perception as reduced. Therefore, shorter vowel durations implicate in a higher application rate of schwa in the perceptual verification. Thus, future researches should collect more data of native speakers of English and verify them both perceptually and acoustically, in order to find if the reduced vowels produced by them are really a schwa or a very short and weak variety of a full vowel.

5. REFERENCES

[1] Chomsky, N., Halle, M. 1968. *The Sound Pattern of English*. New York: Harper & Row Publishers.

[2] Crosswhite, K. 1999. *Vowel Reduction in Optimality Theory*. Los Angeles: UCLA.

[3] Escudero, P., et. al. 2009. A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. *J. Acoust. Soc. of Am.* 126, 1379-1393.

[4] Flanagan, J. et al. 1979. Speech Coding. *IEEE Transactions on Communications* 27(4), 710-737.

[5] Hammond, M. 1997. Vowel Quantity and syllabification in English. *Language* 73, 1-17.

[6] Johnson, K. 2003. *Acoustic and Auditory Phonetics* (2nd ed.). Oxford: Blackwell.

[7] Lieberman, P., Blumstein, S. 1998. *Speech Physiology, Speech Perception, and Acoustic Phonetics.* Cambridge: CUP.

[8] Lindblom, B. 1963. Spectrographic study of vowel reduction. *J. Acoust. Soc. of America* 35, 1773-1781.

[9] Marusso, A. 2003. *Redução Vocálica e Ritmo: Estudo de Caso no Português Brasileiro e no Inglês Britânico*. Dissertation. Belo Horizonte: UFMG.

[10] Reetz, H., Jongman, A. 2009. *Phonetics: Transcription, Production, Acoustics and Perception.* Oxford: Wiley-Blackwell.

[11] Roach, P. 2000. *English Phonetics and Phonology: A Practical Course* (3rd ed.). Cambridge: CUP.

[12] Rosner, B.S., Pickering, J.B. 1994. *Vowel Perception and Production.* Oxford: Oxford University Press.

[13] Yavas, M. 2006. *Applied English Phonology*. Malden: Blackwell Publishers.