LINGUISTIC UNIFORMITY IN THE SPEECH OF BRAZILIAN INTERNAL MIGRANTS IN A DIALECT CONTACT SITUATION

Livia Oushiro

University of Campinas
oushiro@iel.unicamp.br

ABSTRACT

We analyze dialect accommodation of pretonic midvowels /e/ and /o/ in the speech of 32 Northeastern migrants living in the Southeastern state of São Paulo/Brazil. Both vowels undergo lowering in Northeastern dialects (hɛ.ˈlɔ.ʒiw 'watch', hɔ.ˈmã 'pomegranate') but not in the Southeast (cf. he.ˈlɔ.ʒiw, ho.ˈmã). We thus analyzed whether these speakers, in accommodating to the host community's dialect, tend to raise none, one or both vowels. Analyses of 1,916 tokens of /e/ (μF1 = 475Hz, sd = 47) and 1,645 tokens of /o/ (μF1 = 482Hz, sd = 51), extracted from sociolinguistic interviews with the migrants, show that there's a significant but weak correlation between the vowels' height (Spearman's rho=0.38, p<0.05), signaling uniformity in dialect accommodation. Further analyses of the individuals' speech show that half of them accommodated to only one of the vowels, and that accommodation to both vowels depends on speakers' early arrival at the host community.

Keywords: uniformity; dialect accommodation; pretonic vowels; Brazilian Portuguese

1. INTRODUCTION

This paper analyzes linguistic uniformity [3], or linguistic coherence [7], in the speech of Northeastern internal migrants living in the Southeastern state of São Paulo, in Brazil. Earlier works on linguistic coherence have analyzed covariation between variables which are structurally related [6] (e.g. between -s deletion and nominal agreement) and unrelated [10] (e.g. between coda /r/ pronunciation and nasal /e/ diphthongization), among multiple changes in progress [14] (e.g. quotatives *be like/say/do* and intensifiers *so/very/really*), local features [15] (e.g. between verbal -s as in *I likes it* and intensifier *right* as in *He's right huge* in Inuit English) and stereotypes [6] (e.g. verb and nominal agreement in Brazilian Portuguese). Fewer works, in comparison, have examined linguistic coherence in migrants' or immigrants' speech (but see [5]).

In fact, migrants' speech has received less attention than more "prototypical" members of a speech community within language studies in general [4]. But the speech of migrant populations experiencing dialect contact may provide insight into one of the key mechanisms of language change [16] and patterns of dialect acquisition [2, 13].

Northeastern and Southeastern dialects of Brazilian Portuguese differ in several traits, and working-class Northeasterners who have moved to São Paulo in search of better life conditions are likely to be recognized and suffer prejudice due to their speech. One phonetic trait that differentiates Northeasterners and Southeasterners in Brazil is the realization of pretonic midvowels /e/ and /o/, in words such as *relógio* 'watch' and *romã* 'pomegranate': while the pretonic midvowel may undergo lowering in Northeastern dialects when the following syllable contains a low or a nasal vowel (hɛ.ˈlɔ.ʒiw, hɔ.ˈmã; [11]), this type of vowel harmony does not occur in Southeastern dialects (cf. he.ˈlɔ.ʒiw, ho.ˈmã; [1]).

Speakers experiencing dialect contact are, to some extent, expected to accommodate to the host community's speech, acquiring certain linguistic features from their new environment. We thus analyze whether Northeastern migrants in São Paulo, in accommodating to the host community's linguistic patterns, tend to raise their pretonic midvowels and, if so, if this occurs to only one or both midvowels. From the perspective of linguistic uniformity, the expectation is that this process applies simultaneously to both /e/ and /o/.

Strong evidence of uniformity for vowels comes from Labov's [8] seminal work in New York City: pairs of phonetic variables such as (aeh)- and (oh)-raising (as in *bad* and *law*) and (ay) and (aw) (as in *ride* and *loud*) tend to have similar indices for each speaker. This, however, refers to the speech of native speakers.

By contrast, here we found that while there is a significant correlation between pretonic /e/ and /o/'s height in the migrants' speech, further analyses of individuals' speech show that for half of the speakers, convergence with the host community's patterns has occurred for only one of the vowels. We show that acquiring the host community's "target" vowel height for both midvowels is largely dependent on speakers' age of arrival in the new community.

2. CORPUS AND METHODS

We analyzed the speech of 32 migrants from the states of Alagoas and Paraíba, in the Northeastern region of Brazil, who have lived in São Paulo between 1 and 55 years and who moved when they were between 9 and 40 years of age. Migrants were balanced for gender, length of residence, and age of arrival, and were recorded in one-hour-long sociolinguistic interviews by natives of the state of São Paulo. Additionally, the speech of 7 native Paulistanos of similar social characteristics, also recorded in sociolinguistic interviews, was analyzed as a control group, against which accommodation to the host community's pattern was evaluated.

For each of the 39 speakers, we collected 190–200 tokens of pretonic vowels: about 50 tokens of pretonic /e/ and /o/, and about 30 tokens of /i/, /a/ and /u/ (in order to visualize each speaker's full pretonic vowel space), in the phonetic context in which Northeasterners are most likely to produce vowel lowering: words that contain a low or a nasal vowel in the following syllable [11]. Vowels were normalized by the Lobanov method [9] and analyzed in R [12] in a non-parametric correlation test and in mixed-effects linear regression models including Lexical Item as a random effect.

3. RESULTS

Table 1 shows the results of two linear mixed-effects models, one for each vowel (/e/ and /o/), comparing speakers' vowel height by sample (native Paulistanos *vs*. migrants). It shows that, overall, migrants' pretonic vowels are significantly lower (+13.6Hz and +10.0Hz for /e/ and /o/, respectively) than those of native Paulistanos.

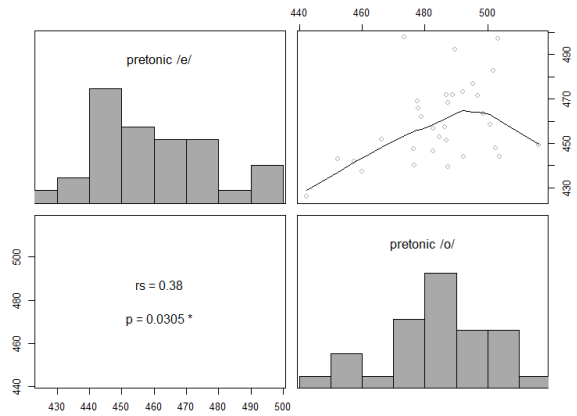
Table 1: Linear mixed-effects models of pretonic /e/ and /o/ height (F1 in Hz).

	Vowel /e/ N = 2,248		Vowel /o/ N = 1,984	
	Est.	*p*	Est.	*p*
Intercept:SP	445.8	***	471.8	***
migrants	13.6	***	10.0	***

***$p < 0.001$.

Figure 1 shows how the two vowels pattern simultaneously for each migrant. It displays the distribution of F1 measurements in the histograms on the top left and bottom right panels, and a scatterplot of migrants' pretonic /e/ and /o/ vowels on the top right panel. On the scatterplot, each dot represents a speaker; the *x*-axis refers to mean F1 measurements of pretonic /o/, and the *y*-axis refers to mean F1 measurements of pretonic /e/. Thus, speakers on the bottom left of that panel are the ones who have mostly converged with the host community's pattern, with lower measurements of F1 (i.e. relatively higher vowels).

Figure 1: Scatterplot of mean pretonic /e/ and /o/ for migrant speakers.



As F1 measurements for both vowels are not normally distributed, a non-parametric correlation test was run (Spearman). There is a significant but weak correlation ($\rho = 0.38$, $p = 0.03$) between the vowel heights. On the scatterplot, we can see that along with speakers whose pretonic /e/ and /o/ are relatively symmetric, there are a few speakers – falling father from the regression line or the mid diagonal – who have favored accommodation to only one of the vowels.

We thus analyzed individual speakers' behavior in linear mixed-effects models, one for each vowel, with the native Paulistanos' sample as the reference level and Lexical Item as a random effect. Tables 2 and 3 show the results for female and male speakers respectively, identified by their pseudonyms.

Table 2 shows that native Paulistanas' average F1 is 449.2 Hz for /e/ and 477.9 Hz for /o/. On the subsequent lines, each speaker's estimate for vowel height is shown as the difference in relation to the reference level "SP", along with its significance. Positive estimates indicate lower vowels than those of native Paulistanas (i.e., closer to the migrants' home state pattern), negative estimates indicate higher vowels than those of native Paulistanas, and estimates close to zero showing no significant differences indicate similar patterns to those of the host community. Since we are interested in which speakers have approximated the São Paulo dialect, we consider both negative estimates and insignificant differences as convergence with the host community's pattern. The last column thus indicates if the speaker has accommodated to none, one or both pretonic midvowels.

Table 2: Linear mixed-effects models of pretonic /e/ and /o/ height (F1 in Hz) for female speakers.

	Vowel /e/ N = 1,149		Vowel /o/ N = 990		
	Est.	*p*	Est.	*p*	Acco.
Intercept:SP	449.2	***	477.9	***	
Fabiana	-9.1		-1.3		both
Gildete	-7.4		-20.3	**	both
Jane	-2.8		4.8		both
Maria	-6.1		-25.6	***	both
Natalia	8.3		8.6		both
Solange	2.7		-11.8		both
Vanessa	-23.2	***	-35.5	***	both
Darlene	-1.2		24.6	**	/e/
Lourdes	9.4		22.9	**	/e/
Marinalva	14.3		20.7	*	/e/
Antonia	16.6	*	-0.1		/o/
Josane	22.7	**	10.9		/o/
Marta	22.6	**	9.0		/o/
Michele	24.1	**	14.1		/o/
Nadia	48.6	***	-4.5		/o/
Raquel	27.6	***	17.6		/o/
Valdenice	19.3	*	9.4		/o/

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

Table 3: Linear mixed-effects models of pretonic /e/ and /o/ height (F1 in Hz) for male speakers.

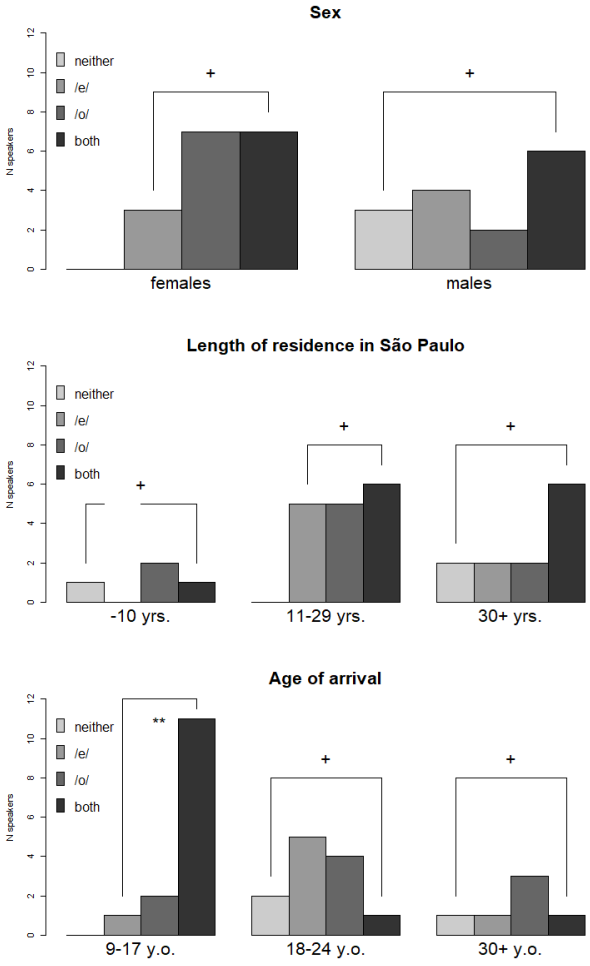
	Vowel /e/ N = 1,099		Vowel /o/ N = 994		
	Est.	*p*	Est.	*p*	Acco.
Intercept:SP	449.9	***	473.2	***	
Aldo	2.8		11.5		both
Geronimo	-12.6	*	-13.3	*	both
JoaoS	6.8		9.5		both
Josue	11.9		5.6		both
Lucas	1.6		13.7		both
Pedro	-2.4		3.2		both
Joao	-5.9		30.5	***	/e/
Marco	-5.8		19.0	*	/e/
PedroC	-0.6		43.1	***	/e/
Valdo	-10.3		14.1	*	/e/
Henrique	42.2	***	16.4		/o/
Josemar	18.9	***	4.3		/o/
Arnaldo	21.5	**	23.6	**	none
Roberto	32.8	***	28.7	***	none
Rodrigo	47.1	***	29.9	**	none

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

On Table 2 we can see that 7 out of the 17 women have accommodated to both vowels, while the others have accommodated to only one of them. A similar pattern arises in the migrant males' speech on Table 3: only 6 out of the 14 speakers have accommodated to both vowels, 6 have accommodated to one of them, and 3 have accommodated to neither.

The question then is who the "coherent" and "incoherent" speakers are and whether there is a clear motivation for their different behavior. In Figure 2, each bar height is the count of the number of speakers accommodating to neither, one or both vowels according to their sociodemographic characteristics: (a) sex; (b) length of residence (less than 10 years, between 11-29 years, or 30+ years); and (c) age of arrival (between 9–17 y.o., 18–24 y.o., or 30+ y.o.). Chi-square tests were run within each group, excluding any empty cells from the frequency tables.

Figure 2: Patterns of accommodation by speakers' social characteristics: sex (top), length of residence in São Paulo (center), and age of arrival (bottom).



+$p > 0.05$; **$p < 0.01$.

As already seen on Tables 2 and 3, both men and women have accommodated to Paulistano vowels to a similar proportion. Although females have apparently accommodated more to vowel /o/ (N = 7) than /e/ (N = 3), and males have accommodated more to vowel /e/ (N = 4) than /o/ (N = 2) in this sample, this may well have happened by chance ($\chi^2 = 1.88(2)$,

p = 0.39 for females and χ^2 = 2.33(3), p = 0.51 for males), as these differences are not significant.

Dividing speakers by length of residence gives us a similar picture. There are no significant differences regarding patterns of accommodation for speakers who have lived at the host community for less than 10 years (χ^2 = 0.50(2), p = 0.78), between 11 and 29 years (χ^2 = 0.12(2), p = 0.90), or more than 30 years (χ^2 = 4.0(3), p = 0.26). Also, there is no difference for those who have been living in São Paulo longer (11 or more years compared to 10 or fewer years; χ^2 = 3.84(2), p = 0.14).

The only clear pattern regarding accommodation to pretonic midvowels relates to speakers' age of arrival: those who arrived at 17 years of age or earlier, i.e., during their late childhood and teenage years, have significantly accommodated to both vowels in comparison to accommodating to only one of them (χ^2 = 13.0(2), p = 0.002). This, however, does not seem to be a gradual trend. For those arriving at 18 or later, there is no significant difference (χ^2 = 3.33(3), p = 0.34 for arrivals between 18 and 24 y.o. and χ^2 = 2.0(3), p = 0.57 for 30+ y.o.).

4. DISCUSSION

The comparison between native Paulistanos and migrants' vowel height in Table 1 shows that migrants have generally not converged with the host community's pretonic mid-vowels, as they maintain significantly lower /e/ and /o/ pretonic vowels.

Individually, however, migrants' speech exhibit great variation. While 13 speakers converged to the host community's pattern regarding both vowel heights, 16 have accommodated to only one of them, and 3 to none. This explains the weak correlation observed in Figure 1 above: only half of the speakers exhibit linguistic uniformity or coherence, in that they have accommodated to neither or to both vowels to a similar degree. The other half has accommodated to only one of the vowels, with no discernible preference for /e/ or /o/.

An examination of these speakers' social characteristics shows that accommodating to both vowels to a similar degree depends fundamentally on the speakers' early arrival at the new community. Migrating later in life, even if a speaker has lived at the new community for a long period, significantly drops the probability that the speaker will accommodate to one or both vowels and, more importantly, that the speaker will maintain linguistic uniformity for pretonic midvowels.

It has been previously argued that linguistic coherence is stronger among speakers whose parents were also born in the same community and among those with low mobility [10]. The overall weak uniformity for pretonic midvowels in our present sample seems to be in accord with that generalization. On the other hand, the fact that half of the 32 mobile speakers have either sustained their native vowels' height or accommodated to both also shows that uniformity is a relevant principle. However, uniformity does not seem to apply to all contexts and speakers. Here we provide evidence that acquiring new parallel phonetic rules, even of a dialect of the same language, depends on relatively early social interaction with speakers of the new community.

The reason why speakers who have accommodated to both vowels are mainly the ones who arrived before the age of 18 is probably due to the social networks they have managed to establish upon arrival [16]. In our sample, the two main reasons for adult migration is searching for better jobs or accompanying a spouse (in the latter case, female speakers whose husbands got a job in São Paulo and are now stay-at-home housewives). Working-class Northeasterners tend to maintain strong ties with other home state migrants in their personal and professional networks. In contrast, those who arrived before the age of 18 have mostly migrated along with their parents and have had the opportunity to attend school with local kids. This type of early and regular social interaction with local peers seems to be the trigger to dialect accommodation. Testing this specific hypothesis falls beyond the scope of the present study, but our results provide sufficient motivating evidence for this new hypothesis to be pursued in a future study.

5. CONCLUSION

This paper has analyzed uniformity or linguistic coherence in the speech of internal migrants in a dialect contact situation. From a starting hypothesis that Northeastern migrants living in the Southeastern state of São Paulo/Brazil would accommodate to two pretonic midvowels whose realization is different in each dialect of Brazilian Portuguese, we found that only half of the speakers are "coherent", either maintaining a lower realization or raising both. Further analyses of speakers' sociodemographic characteristics indicate that accommodation to both vowels is largely dependent on early arrival at the new community, before adulthood.

Future work on uniformity and linguistic coherence should thus also examine other social and internal constraints, in order to explore the limits of its workings. In particular, the role of social networks, language attitudes and identities, albeit difficulties in operationalizing them, should offer insights into the great amount of individual variation and apparent "incoherence" that we find in dialect accommodation.

6. REFERENCES

[1] Cardoso, S. A. M., Mota, J. A., Aguilera, V. A., Aragão, M. S. S., Isquerdo, A. N., Razky, A. M., Wessling, F., Altenhofen, C. 2014. *Atlas Linguístico do Brasil*, vol. 1. Londrina: Eduel.

[2] Chambers, J. K. 1992. Dialect acquisition. *Language* 68(4), 673–705.

[3] Chodroff, E., Wilson, C. 2017. Structure in talker-specific phonetic realization: Covariation of stop consonant VOT in American English. *Journal of Phonetics* 61, 30–47.

[4] Dodsworth, R. 2017. Migration and dialect contact. *Annual Review of Linguistics* 3, 331–346.

[5] Erker, D. 2017. Contact, covariation, and sociolinguistic salience: what Mister Rogers knows about language change. *Penn Working Papers in Linguistics* 23(2), 68–78.

[6] Guy, G. R. 2013. The cognitive coherence of sociolects: how do speakers handle multiple sociolinguistic variables? *Journal of Pragmatics* 52, 63–71.

[7] Guy, G. R., Hinskens, F. 2016. Linguistic coherence: systems, repertoires and speech communities. *Lingua* 172–173, 1–9.

[8] Labov, W. 2006 [1966]. *The social stratification of English in New York City*. Cambridge: Cambridge University Press.

[9] Lobanov, B. M. 1971. Classification of Russian vowels spoken by different speakers. *The Journal of the Acoustical Society of America* 49(2), 606–608.

[10] Oushiro, L. 2016. Social and structural constraints in lectal cohesion. *Lingua* 172–3, 116–130.

[11] Pereira, R. C. 1997. *As vogais médias pretônicas na fala pessoense urbana*. Master's Thesis, João Pessoa, UFPB.

[12] R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

[13] Tagliamonte, S., Molfenter, S. 2007. How'd you get that accent?: Acquiring a second dialect of the same language. *Language in Society* 36(5), 649–675.

[14] Tagliamonte, S., Waters, C. 2011. Co-variation in the speech community: methods for identifying innovators and their repertoires. *Proceedings Methods in Dialectology* 14, London, ON.

[15] Thorburn, J. 2014. *Dialect development in Nain, Nunatsiavut*: emerging English in a Canadian aboriginal community. PhD Dissertation, St. John's (Newfoundland), Memorial University.

[16] Trudgill, P. 1986. *Dialects in contact*. New York: Basil-Blackwell.