

DIPHTHONGIZED (EN) AND THE INDEXATION OF FEMININITY AND PAULISTANITY

RONALD BELINE MENDES¹

RESUMO: A partir de trabalhos sobre produção e avaliação social da pronúncia de /e/ nasal em São Paulo (en), este artigo analisa percepções de femininidade, masculinidade e paulistanidade associadas às variantes dessa variável. Oushiro (2015) mostra que, em palavras como *tempo*, a pronúncia ditongada, relativamente ao monotongo, é favorecida na fala das mulheres, das classes mais altas, com maior grau de escolaridade e quando o estilo (no sentido de Labov 1966) é mais cuidadoso. A área de residência na cidade (central ou periférica) não se correlaciona à variável. Por outro lado, diferentemente do que se verifica para outras variáveis linguísticas, os paulistanos (aqueles que nasceram e cresceram na cidade) não produzem metacomentários com frequência sobre (en) (OUSHIRO 2015). Esses fatos indicam que a pronúncia ditongada de (en) detém prestígio na cidade, a despeito de seu estatuto “abaixo do nível de consciência” (LABOV 2001). Utilizando-se a técnica *matched-guise* (LAMBERT ET AL. 1960, CAMPBELL-KIBLER 2006, LEVON 2014), 8 estímulos auditivos contendo uma única ocorrência de (en) e produzidos por 4 falantes (2 homens e 2 mulheres), foram apresentados a 44 ouvintes, que os julgaram através de várias escalas. No geral, os resultados mostram que a pronúncia ditongada leva a percepções de femininidade, paulistanidade e pertencimento a áreas mais centrais da cidade. Entretanto, há diferenças nesse padrão perceptual de acordo com os falantes. Assim, este artigo discute como fatos de percepção se somam aos de produção e avaliação e avaliação no estudo do significado social das variantes linguísticas.
Palavras-chave: pronúncia ditongada de (en), femininidade, paulistanidade.

RESUMEN: A partir de trabajos sobre producción y evaluación social de la pronunciación de /e/ nasal en San Pablo (en), este artículo analiza percepciones de femineidad, masculinidad y paulistanidad asociadas a las variantes de esa variable. Oushiro (2015) muestra que, en palabras como *tempo* ‘tiempo’, la pronunciación diptongada, en relación al monotongo, se ve favorecida en el habla de las mujeres, de las clases más altas, con mayor grado de escolaridad y cuando el estilo (en el sentido de Labov 1966) es más cuidado. El área de residencia en la ciudad (central o periférica) no se relaciona con la variable. Por otro lado, a diferencia de lo que se verifica en otras variables lingüísticas, los paulistanos (los que nacieron y crecieron en la ciudad) no producen metacomentarios con frecuencia sobre (en) (OUSHIRO 2015). Estos hechos indican que la pronunciación diptongada de (en) tiene prestigio en la ciudad, a pesar de su estatuto “por debajo del nivel de conciencia” (LABOV 2001). Utilizando la técnica *matched-guise* (LAMBERT ET AL. 1960, CAMPBELL-KIBLER 2006, LEVON 2014), se presentaron 8 estímulos auditivos, conteniendo una única ocurrencia de (en) y producidos por 4 hablantes (2 hombres y 2 mujeres), a 44 oyentes que evaluaron esos enunciados a través de varias escalas. En general, los resultados muestran que la pronunciación diptongada lleva a percepciones de femineidad, paulistanidad e pertenecimiento a áreas más centrales de la ciudad. Sin embargo, hay diferencias en ese patrón de acuerdo con los hablantes. Así, este artículo discute cómo los hechos de percepción se suman a los de producción y evaluación en el estudio del significado social de las variantes lingüísticas.

Palabras clave: pronunciación diptongada de (en), femineidad, paulistanidad.

¹ Universidade de São Paulo.

1. INTRODUCTION

In the second edition of *The Social Stratification of English in New York City*, Labov (2006 [1966]:265) adds a thought-provoking comment to the introduction of the third part of his pioneering work:

*“Part III of the book, dealing with “Social evaluation” incorporates results from a series of field experiments that formed an essential part of the methodology. The experimental approaches did not take hold in sociolinguistics in any way comparable to the studies of speech production that formed the basis of Part II. Though each of these experimental methods has had a history of replications and development, the studies that incorporate experiments are few by comparison with those that do not.”*²

Turning to the city of São Paulo, not only are experimental studies focusing on social information extracted from speech rare, but until very recently there also were few production studies about locally relevant linguistic variables. Although the Portuguese spoken in São Paulo has been documented since the 1970s (CASTILHO; PRETI 1986, 1987; PRETI; URBANO 1988, 1990) and analyzed in various works (TARALLO 1983; RODRIGUES 1987; CASTILHO 1990; CALLOU ET AL. 1996; MENDES 2005; COELHO 2006; CASTILHO ET AL. 2006; ILARI 2008; OUSHIRO 2011; among others), Oushiro (2015) can be considered the first comprehensive investigation of the Brazilian variety spoken by those born and raised in the city – the *Paulistanos*. In her work, four linguistic variables are explored both in terms of production and social evaluation: the pronunciation of nasal /e/ in words like *tempo* ‘time’ (as a monophthong or a diphthong); the pronunciation of coda /-r/ (as a tap or a retroflex); noun-phrase number agreement; and subject-verb agreement. In addition, her work includes an analysis of lectal cohesion, by testing the covariation among the four variables, as well as an experiment to test how differently speakers are perceived in their tap /-r/- and retroflex /-r/-guises.

This article contributes to understanding how *Paulistanos* handle language variation in the city by conducting an experimental study about how speakers are perceived when they are listened to in a monophthong or a diphthong-guise of nasal /e/, henceforth referred to as variable (en). The study focuses particularly on perceived gender (masculinity, femininity) and Paulistinity (the quality of sounding like a *Paulistano*), since these are social values that seem to matter in the city, in reference to (en) – based on Oushiro’s (2015) production and evaluation studies. As in Campbell-Kibler (2010:378), the term “perception” here is used to refer to “the processes engaged when people are exposed to external stimuli (...) and extract (social) information from it.”

² Highlights were added to the original.

A more general interest of this article lies in the discussion of how production, evaluation and perception studies are ancillary in the search for explaining socially meaningful language variation. This integration was envisioned by Labov back in the 1960's (in spite of his thoughts – highlighted above – about how experimental work and the development of subjective evaluation tools have not enjoyed the same success as production studies). In celebration of the 50th anniversary of his groundbreaking work, this article is particularly inspired by two of Labov's introductory paragraphs to the section dedicated to subjective evaluations of linguistic variables:

“Most reactions to phonological variables are inarticulate responses below the level of conscious awareness. They occur as part of an overall reaction to many variables. There is no vocabulary of socially meaningful terms with which our informants can evaluate speech for us. We therefore need to proceed not by direct questions, but by eliciting some kind of evaluative behavior that is sensitive enough to reflect the influence of many variables, and is subject to quantitative measurement.

Direct questions are almost useless. Some informants will be ready and willing to answer questions about a certain variable. But the great majority of respondents show no conscious awareness of the variables we have been studying. In the discussion of linguistic attitudes which took place at the end of our interviews, many respondents showed strong opinions about New York City speech in general, but only a few were able to mention specific words, sounds, or phrases which characterized the language of the city or of groups of it. Direct questions will tap the reactions of only a handful of exceptionally articulate middle class speakers.”

2. VARIABLE (EN) IN SÃO PAULO

In her analysis of production of (en), Oushiro (2015:71-88) finds that the diphthongized variant occurs 41% of the time in a pool of 7,235 tokens of word-initial and word-medial stressed nasal /e/, including 50 tokens randomly selected from sociolinguistic interviews with 118 *Paulistanos*, plus all tokens extracted from reading passages solicited from the same informants. Phonological variables, like preceding and following consonant (coded as coronal, labial, dorsal and pause, or as voiceless and voiced), and vowel of the following syllable, do not correlate with (en). The only linguistic variables that predict (en) diphthongization are the morphological word class (with content words – such as *ambiente* ‘environment’ and *pretendo* ‘I intend’ – slightly favoring it, as opposed to adverbials and grammatical words, like the pronoun *a gente* ‘we’ and the preposition *dentro* ‘in’), and morpheme type (stem vs. affix, with diphthongization being slightly favored by the former).

Socially, the production of (en) is influenced by many variables: sex/gender, age group, socioeconomic class, level of education, mobility and style (LABOV 2006 [1966]). Oushiro 2015 interprets this as a case of apparent-time change in progress, in the direction of diphthongization, which is favored by women, younger speakers, higher classes, higher levels of education, lower mobility in the city (that is, by those *Paulistanos* who have lived longer in the same neighborhood), and more careful styles (reading passages and word list). In addition, this change progresses below the level of conscious awareness, given that *Paulistanos* make almost no metalinguistic commentary on the variant, despite the prestige that it seemingly carries.

The idea of prestige, crucial to Labovian sociolinguistics and the concept of style and class stratification, ties in with these quantitative results for the social variables. However, one of the few informants of the study who actually refers clearly to the pronunciation of (en), associates diphthongization with *patricinhas* – girly girls, who belong to higher social classes or affiliate to their values, and who perhaps spend too much time taking care of their hair and nails.³

- Interviewer *e essa frase aqui ó... [...] “você tá **entendendo**”?⁴ você acha esquisito ou normal?*
 ‘what about this sentence... “are you understanding”? do you find it weird or normal?’
- Fernando *“**entendendo**” é meio forçado né*
 ‘“understanding” is a bit forced no’
- Interviewer *forçado? mas você acha que alguém fala assim?*
 ‘forced? but do you think people talk like this?’
- Fernando *acho*
 ‘I do’
- Interviewer *quem?*
 ‘who?’
- Fernando *ah mina principalmente né*
 ‘ah girls mainly right’
- Interviewer *mulher?*
 ‘women’
- Fernando [risos] *principalmente né*
 [laughs] ‘mainly right’
- Interviewer *mas... mulher daqui?*
 ‘but women from here?’
- Fernando *é daqui paulista bem paulista... patizinha⁵ total “cê tá **entendendo**?” né...*
 ‘yes from here *Paulista* very *Paulista*... total *patizinha* “are you understanding?” right...’

³ A good enough visual reference for the São Paulo characterization of *patricinhas* could be the character Cher, in the movie “Clueless” (whose Brazilian version was entitled *Patricinhas de Beverly Hills* ‘Beverly Hills *Patricinhas*’).

⁴ The interviewer started by exaggerating the diphthongized pronunciation of (en) in *entendendo* ‘understanding’. Fernando took the hint and he too exaggerated in this diphthongized pronunciation of (en). His interview can be accessed at <http://projetosp2010.flch.usp.br>

⁵ The term *patizinha* is the diminutive of *paty*, which is a common nickname for *Patricia* in Brazil. Fernando says *patizinha* to mean *patricinha*.

In daily life in São Paulo, such a characterological figure (in the sense of AGHA 2003) is not necessarily prestigious. Moreover, if on the one hand *Paulistanos* do not readily comment on (en), on the other hand people from other Brazilian states express the opinion that (en) diphthongization is “ugly”, “forced”, “funny”, “weird” and “from down South”⁶ – as pointed out by Oushiro (2015) in her assessment of metacommentary offered by interviewees from Minas Gerais (BIELER DA SILVA 2015), Mato Grosso do Sul (VIEIRA 2016) and Maranhão (SANTOS 2015).

Therefore, in light of Labov’s (2006 [1966]) appeal, we need to further investigate what social meanings can be potentially associated with (en) – “potentially” in the sense suggested by Eckert (2008), in her definition of the indexical field of linguistic variants.⁷ Informed by Oushiro’s (2015) work on (en) production and evaluation, the matched-guise experiment that follows was designed in order to attempt to mainly verify if *Paulistanos* do, in fact, associate diphthongized (en) with femininity and Paulistinity.

3. METHOD

Following recent work on sociolinguistic perception – especially Campbell-Kibler (2006 2009, 2011), Phrao *et al.* (2011), Walker *et al.* (2014) and Levon (2014) – stimuli containing one token of (en) were organized in two sets as shown in Table 1.

Table 1: Two sets of (en) stimuli

A	B
Janaína-diphthong	Janaína-monophthong
Lucas-monophthong	Lucas-diphthong
Ariane-monophthong	Ariane-diphthong
Robson-diphthong	Robson-monophthong

The speakers were selected from the Projeto SP2010 corpus.⁸ Lucas’s and Robson’s speech had been previously used by Mendes (2016) in a perception study of noun-phrase (NP) number agreement focusing on perceived gender and

⁶ This last “attribute” was alluded to by an informant interviewed in and native from Maranhão, a Northeastern state.

⁷ Eckert (2008) defines an indexical field as a constellation of ideologically related social meanings that may or may not be activated by a certain speaker-hearer, in a certain situation, when she or he takes a certain stance or projects a certain *persona*.

⁸ Funded by FAPESP (Proc. no. 2011/09278-6), this project collected, transcribed, and made available at (<http://projetsp2010.fflch.usp.br>) is the most recent sample of sociolinguistic interviews with *Paulistanos* in São Paulo to date. All informants were identified by pseudonyms.

sexuality. Lucas was evaluated by the interviewers who worked on the project (graduate and undergraduate students of Linguistics at the University of São Paulo) as the most effeminate-sounding man in the sample, while Robson was evaluated as the most masculine-sounding one. In light of Mendes's (2016) findings – that these men are perceived as less masculine-sounding in their nonstandard plural NP-guise – their speech is also used here, so that comparisons and generalizations can be made, here and in future work. Janaína and Ariane were picked because their ages at the time they were recorded (28 and 36, respectively) are similar to those of Lucas (25) and Robson (31).

The current experiment is part of a larger research project that focuses on perceptions involving (en), NP number agreement and the combination of these two variables. One of the experimental interests of this project is in verifying the differences between listeners' reactions when stimulated by relatively longer stimuli (as in MENDES 2016 – whose audio clips were 20-second long on average and contained 4-7 tokens of plural NPs) versus when stimulated by shorter clips (around 5 seconds). For this reason, the stimuli here contain one single token of (en):

Janaína	<i>Eu acho que aqui em São Paulo você não tem tempo!</i> 'I think that here in São Paulo you don't have time!'
Lucas	<i>Eu moro... super no centro de todos esses bairros.</i> 'I live... very much in the center of all these neighborhoods.'
Ariane	<i>A gente é amiga há muito tempo... agora eu conheço, sou amiga da mulher dele...</i> 'We've been friends for a long time... now I know, I'm friends with his wife...'
Robson	<i>Tem ali uma parte também do centro que é mais chinês...</i> 'There's there a part of downtown too which is more Chinese'

In these tokens, (en) was pronounced originally as a diphthong. Other occurrences of the same words (*tempo* 'time' for Janaína and Ariane; *centro* 'center/downtown' for Lucas and Robson) were also extracted, in which (en) was pronounced as a monophthong. Using Praat (BOERSMA; WEENINK 2014), monophthong (en) was then cut from these tokens and used to replace diphthong (en) in the original tokens, in order to generate the monophthong versions of the same short excerpts. For each of the four speakers, then, two stimuli (the guises) were produced. In an effort to make variable (en) the only "perceptually striking" difference between the guises, all of them were adjusted (in Praat) for intensity, volume and pitch. Each participant in the experiment listened to each speaker in only one guise; in other words, each participant listened to the stimuli in set A or set B. The stimuli were always played for the participants in the order in which they are displayed in Table 1. After listening to each speaker (as many times as she or he wanted), the participant would fill out a perception form, which contained the 9 scales in Figure 1 and other items.⁹

⁹ For a full version of the perception form, see the Appendix.

This girl / This guy sounds:

- not Paulistano at all very Paulistano
- not intelligent at all very intelligent
- not friendly at all very friendly
- not masculine at all very masculine
- not formal at all very formal
- not educated at all very educated
- not extroverted at all very extroverted
- not feminine at all very feminine

What kind of neighborhood does she/he seem to live at?

- more central more peripheral

Figure 1: Scales in the perception form

Listeners were personally contacted by two data collectors – one female and one male Linguistics graduate student¹⁰ at the University of São Paulo – to test whether ratings on various perceptual scales (but mainly those about masculinity/femininity) would vary in a structured way, depending on the data collector’s sex. Guilherme contacted 25 participants (12 women and 13 men); Amanda contacted 19 (8 women and 11 men). This decision was inspired by Hay and Drager’s (2010) work, in which they found that listeners’ perceptions on whether speakers sounded Australian or New Zealander correlated with stuffed toys – a koala bear or a kangaroo versus a kiwi (respectively iconic of these two cultures) – that had been “incidentally” displayed in the room where the experiment was carried.

The great majority of the 44 listeners is in their twenties. The youngest one is 19 years old and the oldest is 65. For this reason, listener’s age was not included in any of the quantitative analyses. This paper concentrates on the scalar responses and the “*patricinha/mauricinho*” checkbox (Cf. Appendix). The *mauricinho* figure is the male correspondent of the girly girl – basically a preppy guy, who makes a continuous effort to appear dashing (perhaps the most iconic element is a cashmere sweater carefully placed over his shoulders). The focus of this checkbox is on the female speakers, but a male version was obviously needed, since the experiment includes male voices as well.

¹⁰ I am indebted and grateful to both Amanda Santana and Guilherme Marchesan, for their participation in this project.

4. FINDINGS

Listener’s ratings on the nine scales were first analyzed in order to check if/how they were correlated with each another. Principal Components Analysis (PCA) in R¹¹ allows for the reduction of these nine scales to four dependent variables (see Table 2): ‘perceived *gender*’ (comprised of the femininity and masculinity scales); ‘perceived *Paulistinity*’ (comprised of the Paulistinity and city area scales); ‘perceived *likeability*’ (comprised of the friendliness and extroversion scales); and ‘perceived *competence*’ (comprised of the intelligence, formality and education scales).¹² As mentioned above, this paper focuses on the first two. Unsurprisingly, the scales for masculinity and femininity are negatively correlated (higher ratings for one correspond to lower ratings for the other). In this case, the PCA indicates that the scales for perceived masculinity and femininity are opposite poles of one underlying component. In other words, as would be expected, the participants did not perceive masculinity and femininity as different or independent scales.

In contrast, the scales for central/peripheral and Paulistinity are positively correlated. While this correlation is not necessarily startling, it is noteworthy that speakers were rated as more *Paulistano*-sounding when they were also rated as sounding like a speaker from a more central area of the city.

Table 2: Factor loading of nine perceptual scales
(Method: Principal Components with varimax rotation)

	‘Gender’	‘Paulistinity’	‘Likeability’	‘Competence’
Paulistano	0.14	0.79	0.09	0.07
Intelligent	0.05	0.05	0.29	0.83
Friendly	-0.15	-0.12	0.86	0.19
Masculine	0.94	0.07	-0.13	0.13
Formal	0.25	0.17	-0.28	0.71
Educated	0.08	0.47	0.00	0.62
Extroverted	-0.23	0.43	0.64	-0.31
Feminine	-0.95	0.03	0.15	-0.12
City-central	-0.12	0.75	-0.11	-0.29
Eigenvalue	1.95	1.82	1.63	1.37
% Variance	22	20	18	15
Cumulative %	22	42	60	75

¹¹ The function used was *principal*, from the package “psych” (REVELLE 2016).

¹² Except for ‘perceived *Paulistinity*’, all of the other three suggested names for the components are based on Levon (2014).

Except for ‘likeability’, all of the other perceived qualities consist of elements (variables) that were part of the production study reported by Oushiro (2015). At this point, then, it is already interesting to note that there is both convergence and divergence between production and perception related to (en). On the one hand, ‘gender’ was verified as the most important social variable in the stratification of (en) in São Paulo (OUSHIRO 2015:79-80) and, here, gender is the first component detected in the perceptual data. On the other hand, while the hypothesis of correlation between (en) and the variable “area of residence” (central vs. peripheral) was not confirmed in production (OUSHIRO 2015:78), here the corresponding perceptual scale (from less central/more peripheral to more central/less peripheral) together with the scale for “Paulistinity” constitutes the component “perceived Paulistinity”.

In the following sections, perceived ‘gender’ and ‘Paulistinity’ are analyzed as dependent variables in regression models that start by including the 44 participants as a random effect, and their demographic characteristics, like sex, education, city zone where they live, their network (which include many, few or no gay friends) as fixed effects. The data collector (Amanda or Guilherme) and the speaker (Janaína, Lucas, Ariane or Robson) were also included as fixed effects.

4.1. Perceived gender

A first look at the data suggests that listeners’ responses seem to differ depending on the guise (monophthong or diphthong) and on the speaker. Figure 2 shows that Robson tends to be rated as very masculine, irrespective of the guise – his ratings tend to be very high in the scale of masculinity and very low in the scale of femininity. As described in the Methods section above, Robson’s voice had already been used in another perception study (MENDES 2016), in which the guises differed in terms of NP number agreement. While Robson was consistently rated as more masculine-sounding than the other males in that experiment, he was also rated as less masculine-sounding in the standard NP-guise (that is, when plural was marked in all NP elements). Here, the (en)-guises do not seem to affect how masculine/feminine Robson sounds.

The patterns for Ariane, Janaína and Lucas are more similar, to the extent that their boxplots are closer to one another. However, while there doesn’t seem to be a difference in terms of the guises for Ariane, Janaína and Lucas tend to be perceived as more feminine-/less masculine-sounding in their diphthong-guise and less feminine-/more masculine-sounding in their monophthong-guise.

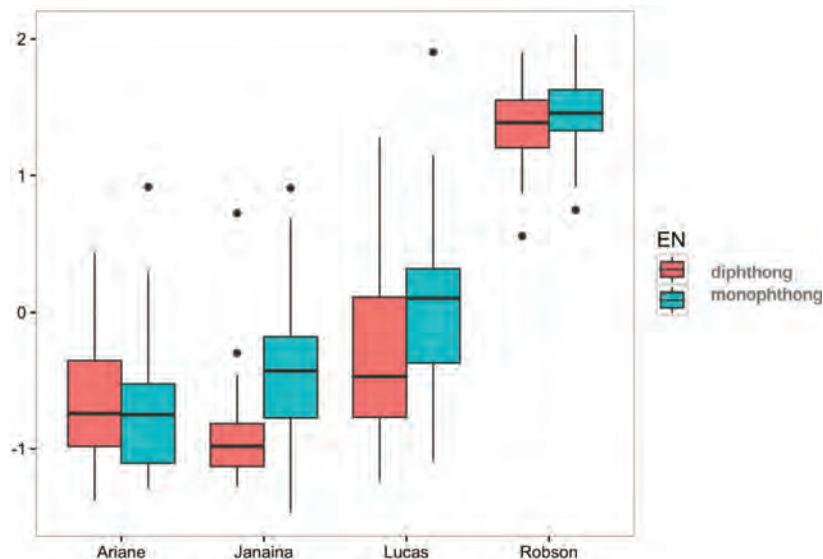


Figure 2: Boxplots for speakers and (en) guises (Principal Component ‘Perceived Gender’)

Initially, various regression models were tested, stepping down from full models that included all of the variables above, plus listener’s sex, level of education, network of friends, zone of residence, data collector (Amanda, Guilherme), and listener as a random effect. None of the listener’s social characteristics play a role in how they rated the stimuli and, in the interest of space, the detailed results for these various models aren’t reported here.

Simpler models (without any of the listener’s social characteristics) were then tested. The first regression model included the speakers, the guises and their interactions, plus listener as a random effect, while a second model included the same variables, but without any interactions. From a third and a fourth model, also without any interactions, the speakers and the guises were respectively excluded. A model selection function (AIC, in the R package MuMIn – BARTON 2016) indicates that the second model is the best, and its results are summarized in Table 3.

Table 3: Linear mixed-model regression results for *perceived gender*

Fixed effects	Estimate	Standard error	<i>t</i> value	<i>p</i> value
(Intercept)	-0.75468	0.08715	-8.660	3.55e-15***
speakerJanaina	0.02095	0.11324	0.185	0.8535
speakerLucas	0.54520	0.11302	4.824	3.10e-06***
speakerRobson	2.04617	0.11373	17.992	< 2e-16***
ENmonophthong	0.20320	0.08053	2.523	0.0125*

Total N: 176. Random effects: Listener (44). Log likelihood: -143.5392

These results show that the speaker is the predictor that most strongly influences the listener’s ratings. However, not only is there an overall guise effect, but the boxplots for Janaína and Lucas above (Figure 2) also lead us to believe that, for them in particular, the guises do matter.

In this analysis, the *intercept* (or reference level) was established as “Ariane” and “ENdiphthong.” Remember that the dependent variable here is the scores (generated by PCA) for the component comprising the masculinity and femininity scales; therefore, the estimated score for Ariane in her diphthong-guise is -0.75468 (that is, low for masculinity). The other values in the estimate column must be read in reference to the intercept. The difference between Janaína and the intercept (Ariane) is rather small (0.02095), and the *p* value over 0.05¹³ is showing that this difference is not statistically significant. Conversely, for Lucas and Robson, the difference between their estimated score values and the *intercept* is significant, as is the difference between the estimated score for ENmonophthong and the reference level (ENdiphthong).

Subsets of data were analyzed for all four speakers separately. Although listener’s sex and data collector (Amanda or Guilherme) do not correlate with perceived gender in the full models mentioned above, the linear models¹⁴ for speaker subsets included these variants, in addition to (en) – since, as previous sections described, they refer to central hypotheses in this perception study. For all speakers, one model included these variables and their interactions, while the other model included no interactions. The results for Janaína (from a model without interactions) and Lucas (from a model with interactions) are in Tables 4 and 5 respectively. For Ariane and Robson, no correlation was verified: Ariane is generally perceived as a feminine-sounding woman, and Robson is generally perceived as a masculine-sounding man – regardless of the (en)-guise, the listener’s sex and whether the researcher collecting data was a woman or a man.

Table 4: Linear regression model results for Janaína’s *perceived gender*

Coefficients	Estimate	Standard error	<i>t</i> value	<i>p</i> value
(Intercept)	-0.66355	0.18900	-3.511	0.00112***
ENmonophthong	0.36012	0.15877	2.268	0.02879*
SexMasc	-0.03245	0.16022	-0.203	0.84050
DatCollectGuilherme	-0.23283	0.15726	-1.481	0.14655
Total N: 44. Log likelihood: -30.45921				

¹³ The conventionalized limit of 5% chance of observing the same pattern in case the null hypothesis is true.

¹⁴ The listeners were not included as a random effect here.

Table 5: Linear regression model results for Lucas's *perceived gender*

Coefficients	Estimate	Standard error	<i>t</i> value	<i>p</i> value
(Intercept)	-0.2331	0.2754	-0.846	0.4030
ENmonophthong	1.4722	0.5509	2.673	0.0112*
SexMasc	-0.3164	0.3895	-0.812	0.4219
DatCollectGuilherme	0.1218	0.4085	0.298	0.7673
ENmono:DataCollectorGuilherme	-1.3383	0.7129	-1.877	0.0686•
SexMasc:DatCollectorGuilherme	0.2532	0.5548	0.456	0.6509
ENmono:SexM:DatColGuilherme	0.9585	0.8893	1.078	0.2883
Total N: 44. Log likelihood: -40.70343				

Both Janaína and Lucas are perceived as more masculine/less feminine in their monophthong-guise. For Lucas, however, the interaction between the guises and the data collector approaches significance (p value = 0.0686) when the latter is another man (Guilherme) and when Lucas is listened to in his monophthong-guise. It is still not safe to reject the null hypothesis here, but nothing similar is observed for Janaína, whose perceived gender depends only on the guise.

Due to the way the experiment was designed, it is not clear why (en) matters for Janaína's and Lucas's perceived gender, but not for the other two speakers'. As noted above, every participant in the experiment listened to the speakers in the same order (Janaína, Lucas, Ariane, Robson). One could argue that some sort of priming could be happening between the female voices, on the one hand, and between the males voices, on the other. One way to solve this problem would be to present the four stimuli to the listeners in all possible orders, which would necessarily require a much greater number of listeners. However, the present results do suggest that there is a correlation between perceived gender and the variants of (en), although such correlation is not essential for all speakers.

Two reflections are in order here. First, considering that Lucas had previously been picked out of a pool of 30 *Paulistano* men as sounding particularly effeminate, while Robson was picked out as sounding particularly masculine (MENDES 2016), one could argue that diphthongized (en) “adds” to Lucas's feminine-soundingness, but doesn't change the perception that Robson sounds masculine. This consideration parallels the hypotheses tested (and to a certain extent confirmed) by Campbell-Kibler (2008, 2009), Pharaoh *et al.* (2011) and Levon (2014). Even though their specific methods and variables differ, these authors essentially discuss a similar idea: that the social meaning of variation is, itself, variable, and is context-, listener- and speaker-dependent. Unlike Lucas and Robson, previous evaluations on Ariane and Janaína were not assessed (for femininity or any other aspect), so nothing can be purported here to potentially explain the different patterns shown in Figure 1 and attested by the regression models. Still, the results presented here serve as further evidence that different speakers may be perceived differently even when the same linguistic variable is examined experimentally.

Secondly, and consequently, these results constitute yet another piece of evidence in favor of the dynamism of socially meaningful variation, which in turn serves as both an argument for and a product of agency (ECKERT 2012). Even though there's a strong correlation between (en) and sex/gender in production in São Paulo, with the diphthongized form being strongly favored by women and disfavored by men, it doesn't follow that every man in São Paulo will always sound more feminine because of that variant, nor that every woman will necessarily sound less so if she pronounces (en) as a monophthong. This is by no means a new thought, but contributes to our understanding in the study of how we manage language variation in our lives.

4.2. Perceived Paulistanity

The second component identified in the data through PCA is composed of the scales for Paulistanity and peripheral-central São Paulo, which are positively correlated: listeners who tended to rate the speakers as more *Paulistano*-sounding also tended to rate them as residents of more central areas of the city; those who tended to rate the speakers as less *Paulistano*-sounding, perceived them as *Paulistanos* from more peripheral areas of the city. As noted earlier, the variable “center vs. periphery” does not correlate to variable (en) in production (OUSHIRO 2015).

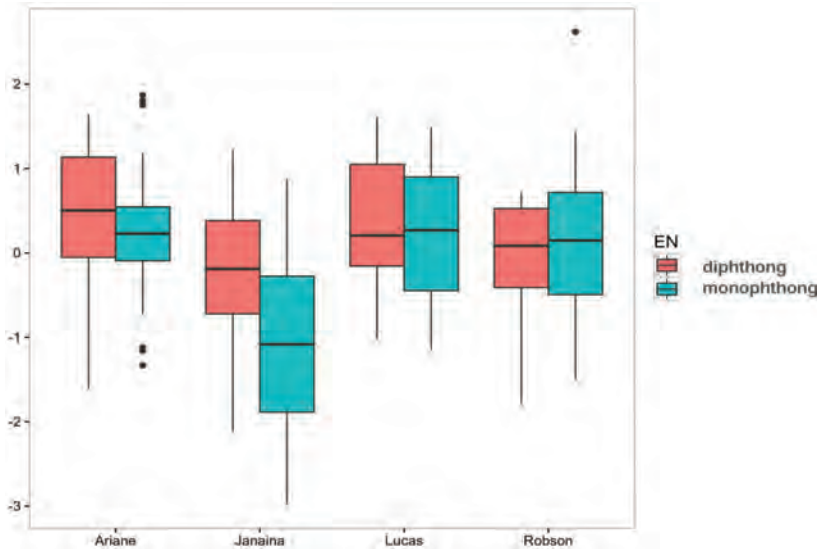


Figure 3: Boxplots for speakers and (en) guises (Principal Component ‘Perceived Paulistanity’)

Figure 3 indicates an apparent divide between the female and male speakers: the component shows less variation for Lucas and Robson, in comparison to Ariane and Janaina, especially the Janaina.

For this component, similar regression models were tested, stepping down from fuller to simpler models. Once again, no social characteristics related to the listeners were attested as having an effect on the dependent variable. The model selection function AIC (in the R package MuMIn) indicated that the best model was the one that included speaker, (en) and listener (random effect), without any interactions. See the results in Table 6.

Table 6: Linear mixed-model regression results for *perceived Paulistinity*

Fixed effects	Estimate	Standard error	<i>t</i> value	<i>p</i> value
(Intercept)	0.39168	0.15118	2.591	0.0104*
speakerJanaína	-0.92390	0.19203	-4.811	4.15e-06***
speakerLucas	0.01552	0.19166	0.081	0.9356
speakerRobson	-0.18643	0.19286	-0.967	0.3355
ENmonophthong	-0.23596	0.13657	-1.728	0.0864•
Total N: 176. Random effects: Listener (44). Log likelihood: -238.1483				

Although there is a significance-approaching effect of (en), these results confirm that Janaína is the speaker whose perceived Paulistinity is impacting the data. In a separate analysis of her data subset, a linear model that includes (en) and the listener’s level of education as predictors (without interactions) attests that she is perceived as someone who sounds less *Paulistano* in her monophthong-guise.

Table 7: Linear regression model results for Janaína’s *perceived Paulistinity*

Coefficients	Estimate	Standard error	<i>t</i> value	<i>p</i> value
(Intercept)	-1.2616	0.7769	-1.624	0.1122
ENmonophthong	-0.7961	0.3107	-2.562	0.0143*
EducationHighSchool	0.6498	0.7953	0.817	0.4188
EducationCollege	1.2025	0.7502	1.603	0.1168
Total N: 44. Log likelihood: -60.64021				

Although the listener’s level of education does not predict how Janaína is perceived for Paulistinity, it does play a role for both Ariane and Robson. Linear models for these two speakers’ data subsets reveal an interaction between (en) and level of education (see Tables 8 and 9). This is interesting to the extent that the listener’s own area of residence (central or peripheral) is not a predictor of how they perceive these four speakers in terms of Paulistinity/city area.¹⁵ All in all, despite the fact that Janaína is perceived as more or less *Paulistano*-sounding

¹⁵ In all models that included listener’s city area of residence (central or peripheral) and/or city zone, with or without interactions, no correlation was found between these and the dependent variable.

depending only on the (en)-guise, the results for Ariane and Robson allow for the interpretation of a rather complex scenario, where there is both a match and a mismatch between production and perception. We have a mismatch when the differentiation between central and peripheral São Paulo – which can be very striking both in social terms and in sociolinguistic patterns – does not influence (en) in production, but does matter in perception (remember that perceived *Paulistanity* here consists of two scales, positively correlated: *Paulistano*-soundingness and centrality-peripherality). On the other hand, we have a match when diphthongized (en) is favored by higher levels of education in production (Oushiuro 2015) and, here, we see listener’s responses varying according to their level of education – at least for Ariane and Robson.

Table 8: Linear regression model results for Ariane’s *perceived Paulistanity*

Coefficients	Estimate	Standard error	<i>t</i> value	<i>p</i> value
(Intercept)	-1.327	0.9012	-1.501	0.1416
ENmonophthong	3.1524	1.2745	2.473	0.0180*
EducationHighSchool	1.4034	1.0076	1.393	0.1718
EducationCollege	1.8083	0.9235	1.958	0.0576
ENmono:EducationHighSchool	-2.9860	1.4010	-2.131	0.0396*
ENmono:EducationCollege	-3.4675	1.3163	-2.634	0.0121*
Total N: 44. Log likelihood: -54.63115				

Table 9: Linear regression model results for Robson’s *perceived Paulistanity*

Coefficients	Estimate	Standard error	<i>t</i> value	<i>p</i> value
(Intercept)	0.7190	0.8112	0.886	0.3810
ENmonophthong	-1.9576	1.1472	-1.706	0.0961
EducationHighSchool	-1.1816	08762	-1.349	0.1855
EducationCollege	-0.5732	0.8473	-0.676	0.5028
ENmono:EducationHighSchool	2.7772	1.2610	2.202	0.0338*
ENmono:EducationCollege	1.9622	1.1863	1.654	0.1063
Total N: 44. Log likelihood: -50.00094				

Finally, Lucas is the only speaker for whom no effect was found in the analysis of perceived Paulistanity. Similarly to what was discussed in the previous section, a potential explanation for this result could be related to masking effects: (en) matters for Lucas’s perceived femininity, but perceptions of Paulistanity are possibly blocked, in relation to the same linguistic variable. Further research would be needed to test this, in the specific case of (en) in São Paulo, but effects of this sort have been tested and discussed in previous work (PHARAO ET AL. 2011, LEVON 2014, among others).

4.3. *Patricinha/Mauricinho*

By relying on Oushiro's (2015) interviewee (Fernando) who evoked the image of a girly girl (*patricinha*) when asked what he thought about diphthongized (en), one could expect that, in this matched-guise experiment, the diphthong-guise would stimulate the listeners to recall the same image for the female speakers. However, the listeners did not check the corresponding box often, as shown in Table 10.

Table 10: Number of times that the *patricinha/mauricinho* box was checked by the 44 listeners, per speaker

	N checked <i>patricinha/mauricinho</i> box	%
Ariane	11	25
Janaína	4	9
Lucas	5	11
Robson	1	2

A generalized mixed-effects model for this data, which includes speaker, (en), listener's sex, and data collector as fixed effects, plus listener as a random effect, indicates that checking the *patricinha/mauricinho* box is not correlated to any of these variables. The only significant effect found is for speaker Robson – which would be expected, since he was perceived as *mauricinho* by only one single participant.

Does this mean that Fernando was wrong when he associated diphthongized (en) with *patricinhas*? Probably not. Such an association was made and there's no reason to deny it. What the data here does show is that (en) does not *necessarily* lead *Paulistanos* to perceptions of a girly girl or a preppy man. In fact, as in Labov's (2006 [1966]) view, Fernando was one of only two (in a pool of 118 informants) who were able to come up with such a specific social figure in association with diphthongized (en). So the fact that (en) does not have an effect in how listeners checked the *patricinha/mauricinho* box is actually consistent with the fact that so few of Oushiro's informants made such an association. Further research could attempt to check how (en), along with other variables, is used as a resource in the construction of styles and urban identities in São Paulo.

5. CONCLUSION

Results of production, evaluation and perception clearly do not always match and, when they do, it may happen that they match only partially – for example, not for all speakers in an experiment such as the one reported here. Therefore, correlations between social variables and dependent linguistic variants should not be taken as direct indications of social identities, be they group- or individual-based. Conversely, although hypotheses about perceptual correlations can be inspired by work on

production and social evaluation, the researcher should not be too surprised if such hypotheses are not confirmed – since the interpretation of results of production studies have been traditionally centered around the concepts of prestige and class (especially given the most traditional methods of population sampling and data collection). Prestige, as with every possible social meaning, is itself variable – as decades of sociolinguistic studies since Labov 1966 have taught us. If we want to have a more comprehensive understanding of what social meanings we operate with in our lives and in what way, and in relation to which linguistic variables, it appears that we have no choice but to follow a direction that Labov pointed out to us long ago: we have to continue to try to elicit the speaker-hearer evaluative behavior, “that is sensitive enough to reflect the influence of many variables” (LABOV 2006 [1966]:265).

REFERENCES

- AGHA, Asif (2003) The social life of cultural value. *Language & Communication* 23: 231–273.
- BARTOŃ, Kamil (2016) *MuMin*: Multi-Model Inference. R package version 1.15.6. Available at <https://CRAN.R-project.org/package=MumIn>.
- BIELER DA SILVA, Mariane (2015) *Entre duas metrópoles: (-R) em Itanhandu*. Master’s Thesis. Universidade de São Paulo.
- BOERSMA, Paul; WEENINK, David (2014) “Praat: doing phonetics by computer”, 2014. Accessed in October 29, 2016. Available at <http://www.fon.hum.uva.nl/praat/>.
- CAMPBELL-KIBLER, Kathryn (2006) *Listener perceptions of sociolinguistic variables: the case of (ing)*. Ph.D dissertation. Stanford University.
- CAMPBELL-KIBLER, Kathryn (2008) I’ll be the judge of that: Diversity in social perceptions of (ING). *Language in Society* 37: 637–59.
- CAMPBELL-KIBLER, Kathryn (2009) The nature of sociolinguistic perception. *Language Variation and Change* 21: 135–56.
- CAMPBELL-KIBLER, Kathryn (2010) Sociolinguistics and Perception. *Language and Linguistics Compass* 4/6: 377–289.
- CAMPBELL-KIBLER, Kathryn. Intersecting variables and perceived sexual orientation in men. *American Speech* 86:52–68, 2011.
- CALLOU, Dinah; MORAES, João; LEITE, Yonne (1996) Variação e diferenciação dialetal: a pronúncia do /r/ no português do Brasil. In: *Gramática do português falado, vol. VI*. Campinas: Editora da Unicamp.
- CASTILHO, Ataliba T. (ed.) (1990) *Gramática do português falado*. Campinas: Editora da Unicamp.
- CASTILHO, Ataliba T.; PRETI, Dino (eds.) (1986) *A linguagem falada culta na cidade de São Paulo: materiais para seu estudo*, vol. I – Elocuções Formais. São Paulo: T.A. Queiroz.
- CASTILHO, Ataliba T.; PRETI, Dino (eds.) (1987) *A linguagem falada culta na cidade de São Paulo: materiais para seu estudo*, vol. II – Diálogos entre dois informantes. São Paulo: T.A. Queiroz/FAPESP.

CASTILHO, Ataliba T.; JUBRAN, Clélia Cândida A. S.; KOCH, Ingedore G. V. (eds.) (2006) *Gramática do português culto falado no Brasil*. Campinas: Editora da Unicamp.

COELHO, Rafael F. *É nós na fita! Duas variáveis linguísticas numa vizinhança da periferia paulistana. O pronome de primeira pessoa do plural e a marcação de plural no verbo*. Master's Thesis. São Paulo: FFLCH-USP, 2006.

ECKERT, Penelope (2008) Variation and the indexical field. *Journal of Sociolinguistics*, vol. 12(4): 453–476.

ECKERT, Penelope (2012) Three waves of variation study: the emergence of meaning in the study of variation. *Annual Review of Anthropology* 41: 87–100.

HAY, Jennifer; DRAGER, Katie (2010) Stuffed toys and speech perception. *Linguistics*, vol. 48(4): 865–892.

ILARI, Rodolfo (ed.) (2008) *Gramática do português falado culto no Brasil*. Campinas: Editora da Unicamp.

LABOV, William (2001) *Principles of linguistic change: social factors*. Oxford & Cambridge: Blackwell.

LABOV, William (2006 [1966]) *The social stratification of English in New York City*. São Paulo: Cambridge University Press.

LAMBERT, Wallace E.; HODGSON, R. C.; GARDNER, R. C; FILLENBAUM, S. (1960) Evaluational reactions to spoken languages. *Journal of Abnormal and Social Psychology* 60(1): 44–51.

LEVON, Erez (2014) Categories, stereotypes, and the linguistic perception of sexuality. *Language in Society* 43: 539–566.

MENDES, Ronald Beline (2005) *Ter + participio e Estar + gerúndio: aspecto verbal e variação no português do Brasil*. PhD Dissertation. Campinas: IEL/Unicamp.

MENDES, Ronald Beline (2016) Nonstandard noun phrase agreement as an index of masculinity. In: LEVON, Erez; MENDES, Ronald Beline (eds.). *Language, sexuality and power: studies in intersectional sociolinguistics*. Oxford: Oxford University Press.

OUSHIRO, Livia (2011) *Uma análise variacionista para as Interrogativas-Q*. Master's thesis. São Paulo: FFLCH-USP.

OUSHIRO, Livia (2015) *Identidade na pluralidade: Avaliação, produção e percepção linguística na cidade de São Paulo*. PhD Dissertation. São Paulo: FFLCH-USP.

PHARAO, Nicolai; MAEGAARD, Marie; MØLLER, Janus S.; KRISTIANSEN, Tore (2014) Indexical meanings of [s+] among Copenhagen youth: Social perception of a phonetic variant in different prosodic contexts. *Language in Society* 43: 1–31.

PRETI, Dino; URBANO, Hudinilson (eds.) (1988) *A linguagem falada culta na cidade de São Paulo: materiais para seu estudo*, vol. III: Entrevistas. São Paulo: T. A. Queiroz/FAPESP.

PRETI, Dino; URBANO, Hudinilson (eds.) (1990) *A linguagem falada culta na cidade de São Paulo: materiais para seu estudo*, vol. IV: Estudos. São Paulo: T.A. Queiroz/FAPESP.

RODRIGUES, Angela C. S. (1987) *A concordância verbal no português popular em São Paulo*. PhD Dissertation. São Paulo: FFLCH-USP.

- REVELLE, W. (2016) *psych*: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA, Version = 1.6.9. Available at <https://CRAN.R-project.org/package=psych>.
- SANTOS, Wendel Silva dos (2015) *A morfologia do indicativo na expressão do modo subjuntivo em São Paulo e São Luís*. Master's Thesis. São Paulo: FFLCH-USP.
- TARALLO, Fernando Luiz (1983) *Relativization strategies in Brazilian Portuguese*. Ph.D. Dissertation. Philadelphia: University of Pennsylvania.
- VIEIRA, Marília (2016) *Ai, daí e então em Campo Grande e São Paulo: Análise sociofuncionalista no domínio da causalidade*. Ph.D. Dissertation. São Paulo: FFLCH-USP.
- WALKER, Abby; GARCÍA, Christina; CORTÉS, Yomi; CAMPBELL-KIBLER, Kathryn (2014) Comparing social meanings across listener and speaker groups: The indexical field of Spanish /s/. *Language Variation and Change* 26: 169–189.

APPENDIX

Esse cara/ essa moça parece:

nada paulistano/a	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	muito paulistano/a
nada inteligente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	muito inteligente
nada amigável	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	muito amigável
nada masculino/a	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	muito masculino/a
nada formal	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	muito formal
nada escolarizado/a	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	muito escolarizado/a
nada extrovertido/a	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	muito extrovertido/a
nada feminino/a	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	muito feminino/a

Em que bairro essa pessoa parece morar?

mais periférico mais central

Em que zona da cidade você acha que ele/ela mora?

Norte Sul Leste Oeste Centro

Qual você acha que é a faixa etária dele/dela?

adolescência 20 e poucos anos 30 e poucos anos 40 ou mais

Qual você acha que é o nível de escolaridade dele/dela?

Fundamental 1 Fundamental 2 Médio Superior

A que classe econômica você acha que ele/ela pertence?

baixa média baixa média média alta alta

Quais dessas características descrevem essa pessoa, na sua opinião?

alto/a	atraente	sincero/a	gay/lésbica
tímido/a	trabalhador/a	caipira	metido/a
descolado	mauricinho/patricinha	articulado/a	mimado/a