

Perception of height and categorization of Brazilian Portuguese front vowels*

Percepção da altura e categorização das vogais anteriores do português brasileiro

Daniel Márcio Rodrigues SILVA (UFMG)

Rui ROTHE-NEVES (UFMG)

ABSTRACT

Cross-linguistic typological observations and theoretical models in phonology suggest that certain speech sound distinctions are more complex than others. One such example is the opposition between mid-high and mid-low vowels, usually thought to be more complex than the opposition between high and mid vowels. The present study provides experimental evidence on speech sound perception which supports this notion. Native Brazilian Portuguese speakers performed vowel classification tasks involving either the distinction between the front high mid /e/ and the front high /i/, or the distinction between the front high mid /e/ and the front low mid /ɛ/ vowel. Measures of response time and discriminability (d') at the vowel category boundaries were obtained. Participants showed significantly slower responses and lower d' values in the “e-ɛ” as compared to the “i-e” classification task. Results indicate that perceptually distinguishing /e/ from /i/ requires more processing time and resources, and involves more complex information than distinguishing /e/ from /i/.

Key-words: *speech perception; categorization; vowel height; sound oppositions.*

*. **Acknowledgments** - The first author is supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) - Brazil. The second author is supported with a research grant by CNPq - Brazil.

RESUMO

Observações tipológicas interlinguísticas e modelos teóricos em fonologia sugerem que certas distinções fônicas são mais complexas que outras. Um exemplo é o caso da oposição entre vogais médias altas e médias baixas, usualmente considerada mais complexa que a oposição entre vogais médias e altas. O estudo aqui apresentado fornece evidências experimentais que apontam na mesma direção. Falantes nativos do português brasileiro realizaram duas tarefas de classificação de vogais, um das quais envolvendo a distinção entre /i/ e /e/; a outra, entre /e/ e /ɛ/. Medidas de tempo de resposta e de discriminabilidade (d') foram obtidas. Respostas mais lentas e valores inferiores de d' foram observados na tarefa de classificação "e-ε". Os resultados indicam que, em comparação à distinção entre /i/ e /e/, a que se estabelece entre /e/ e /i/ requer mais tempo e recursos de processamento e envolve informação mais complexa.

Palavras-chave: percepção da fala; categorização; altura vocálica; oposições sonoras.

Introduction

A foundational observation in phonology is that the sound systems of languages are systems of oppositions. Each vowel and consonant in any given language is defined by its opposition to all others in terms of certain features that are taken as relevant in that particular sound system. In German, for example, there is an opposition between /h/ and /ʁ/, as shown by minimal pairs such as *hose* 'pants' and *rose* 'rose'. No such opposition occurs in Portuguese, as the sounds [h] and [ʁ] are functionally equivalent – using one or the other cannot result in different words/morphemes being perceived by a native listener. Portuguese speakers identify both ['kahʊ] and ['kaʁʊ] as exemplars of the word *carro* 'car'. This example illustrates the fact that speech sounds form functional equivalence classes: sound categories whose elements are treated as functionally equivalent, i.e., having the same value in the language system.

It seems not to be the case that sound systems of languages are formed by collections of oppositions among sound categories drawn from a set of equally likely or equally complex possibilities.

Some sounds (and sound oppositions) are very common; others are rare (Maddieson, 1984). Furthermore, sound inventories are seen as comprising basic elements that are found very frequently across and within languages, and less basic elements through which they can be extended (Maddieson, 1999).

As for vowels, the most common across human languages are /a/, /i/, and /u/ (according to the UPSID database; Maddieson, 1984). Oppositions between high-mid and low-mid vowels are less common than oppositions between mid and high vowels. The most widespread pattern across languages is represented by five-vowel systems comprising a high front, a high back, a mid front, a mid back, and a low vowel (Diehl, 2008; Schwartz, Boë, Vallée, & Abry, 1997a, 1997b). This is the case, for example, of the Spanish vowel system (/i, u, e, o, a/), in which no distinction is made between high-mid and low-mid vowels. The inclusion of this distinction (both for front and mid vowels) results in a seven-vowel system, also relatively common but considerably less widespread than the former. This is the case of the Portuguese vowel system, in which the opposition between high-mid /e, o/ and low-mid vowels /ɛ, ɔ/ – as attested by minimal pairs such as *s[e]de* ‘thirst’ versus *s[ɛ]de* ‘headquarters’; *c[o]rte* ‘court’ versus *c[ɔ]rte* ‘cut’ – is functional only in stressed syllables (Bisol, 2003; Mateus & d’Andrade, 2000; Wetzels, 2011).

In Brazilian Portuguese, the neutralization of the high-mid/low-mid opposition reduces the seven vowels to five in pre-stressed syllables, and the neutralization of the mid/high opposition in word-final post-stressed syllables results in a further reduction to three vowels which are realized as [ɪ], [ʊ], and [ɐ] in most Brazilian Portuguese dialects. Here one notes a progression in which the oppositions /e/-/ɛ/ and /o/-/ɔ/ (that occur less frequently across languages) are the first ones to be neutralized, followed by the oppositions /e/-/i/ and /o/-/u/ (Bisol, 2003; Mateus & d’Andrade, 2000; Wetzels, 2011). Interestingly, some free variation between high-mid and low-mid vowels is observed even in stressed syllables, as in the cases of *p[o]ça* ~ *p[ɔ]ça* ‘puddle’ and *[e]xtra* ~ *[ɛ]xtra* ‘extra’. Therefore, cross- and intra-linguistic findings suggest that certain oppositions are less stable and more complex than others, such as, for example, /e/-/ɛ/ as compared to /i/-/e/ (for related examples in Italian and Catalan, see Krämer, 2009; Wheeler, 2005).

The above observations raise issues about whether and how such differences are reflected in the processing and storage of information related to sound categories and the distinctions that define them. To the extent that phonology is about mental entities and operations, it can provide some (rather indirect) clues. Some phonological analyses of Portuguese are consistent with the idea that certain oppositions involve more complex representations than others. Based on Contrastive Hierarchy Theory (Dresher, 2009), Lee (2010) suggests that Brazilian Portuguese vowels are specified according to a hierarchy of distinctive features (see Figure 1) in which the feature [low] separates /a/ from the remaining six vowels and occupies the highest position, followed by [back], [high], and [ATR]. The latter three features distinguish, respectively, back from front, high from mid, and high-mid from low-mid vowels. Thus, the oppositions /e/-/ɛ/ and /o/-/ɔ/ are the most complex since they require that the whole hierarchy be traversed.

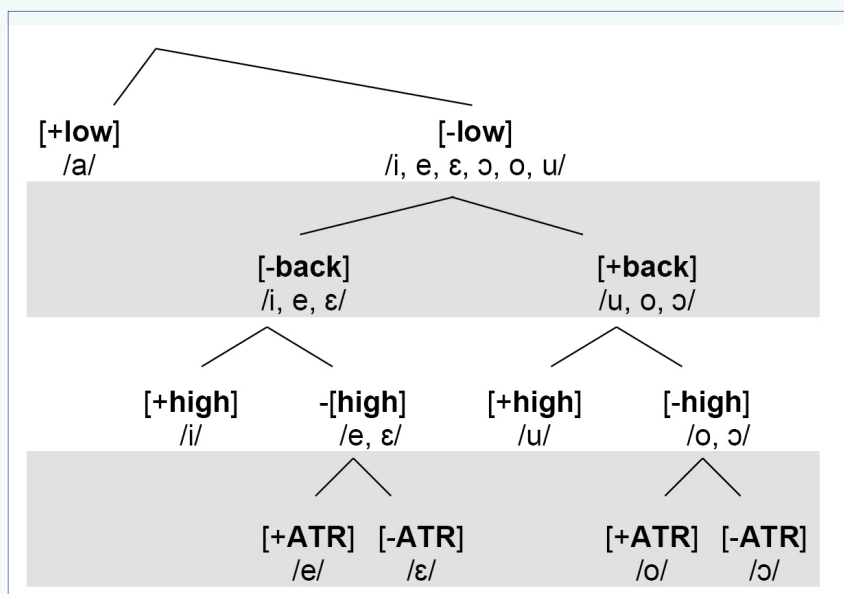


Figure 1 – Contrastive hierarchy for the Brazilian Portuguese vowel system as proposed by Lee (2010).

Another example is found in the analysis presented by Nevins (2012), based on Element Theory, according to which vowels can be described as combinations of three elements: |A|, |I| and |U| (Harris,

1994; Harris & Lindsey, 2000). These elements correspond to the “corner vowels” /a/, /i/ and /u/, considered in the theory as atomic primes. The remaining vowels are derived from combinations among elements. Nevins (2012) describes the mid vowels of Brazilian Portuguese as resulting from the combinations |A|+|I| (for /e/ and /ɛ/) and |A|+|U| (for /o/ and /ɔ/). The distinction between high-mid /e, o/ and low-mid /ɛ, ɔ/ is made possible by allowing vowels to differ in the relative prominence of each element in the representation. Accordingly, the vowels /e, ɛ, ɔ, o/ can be represented as |IA|, |IA|, |UA| and |UA| – where the underline indicates the dominant (head) element¹. Again, the most complex oppositions are those between high-mid and low-mid vowels.

Data on vowel acquisition in Brazilian Portuguese as a native language (Bonilha, 2004) provide further evidence that, from the point of view of speech production, the distinctions between high-mid and low-mid vowels are the last to be acquired – something also predicted in the model of Fikkert (2005) for European Portuguese. In fact, the data provided by Bonilha (2004) are consistent with the feature hierarchy proposed by Lee (2010).

Concerning speech perception, Bosch and Sebastián-Gallés (2003) provide interesting results from Spanish, with its five-vowel system, and Catalan. The number of vowels varies across Catalan dialects, but they generally distinguish /e/ from /ɛ/ and /o/ from /ɔ/ – a seven-vowel inventory is the most widespread (Recasens & Espinosa, 2009; Wheeler, 2005). Four- and eight-month-old infants from Spanish monolingual, Catalan monolingual, and bilingual environments were tested for their ability to discriminate between /e/ and /ɛ/ (a distinction that is absent in Spanish). While four-month-olds from the three linguistic environments performed equally well, Spanish monolingual and bilingual eight-month-olds did not succeed. These results indicate that general perceptual abilities that are present early in development include the ability to discriminate /e/ from /ɛ/, which declines between four and eight months of age unless the linguistic environment establishes a clear separation between the two categories. Results of an additional

1. In the analysis proposed by Nevins (2012), Brazilian Portuguese dialects differ in whether low-mid or high-mid vowels are represented as headed. Further details are beyond the scope of this paper.

experiment indicate that infants raised in the bilingual environment regain the sensitivity to the /e/-/ɛ/ distinction between eight and twelve months of age (Bosch & Sebastián-Gallés, 2003).

The development of language-specific perception of speech sounds involves improvement in sensitivity to distinctions that are used in the language of exposure as well as decline in sensitivity to distinctions that are not (Kuhl, 2004; Kuhl et al., 2008; Werker & Tees, 2005). It has been proposed that, thanks to general auditory processing mechanisms, young infants are sensitive to all distinctions in the world's languages (Kuhl, 2004; Kuhl et al., 2008; but see Mazuka, Hasegawa, & Tsuji, 2014). One example is the ability to distinguish /ɛ/ from /e/, the maintenance of which depends on linguistic experience – as shown by Bosch and Sebastián-Gallés (2003). However, while it may be true that infants are initially able to distinguish all sounds of human speech, evidence indicates that certain sounds are special, such as the corner vowels /a/, /u/ and /i/, which are proposed to function as salient and stable reference points in the vowel space during the acquisition of vowel systems (Polka & Bohn, 2003, 2011).

The present study aims to provide experimental evidence on whether theoretically expected differences in the degree of complexity among vowel oppositions in Brazilian Portuguese are reflected in the way perceptual information related to vowel categories is processed and stored in the speaker's memory. Speakers performed two classification tasks involving either the distinction between /e/ and /i/ or the distinction between /e/ and /ɛ/, which is thought to be more complex, and hence less stable. If the latter involves more information and/or requires more complex operations than the former, we should expect slower and less precise responses in the task that requires distinguishing the categories /e/ and /ɛ/. Therefore, the two tasks were compared in order to test the following two hypotheses: 1) a sharper distinction between categories will be observed in the “i-e” classification task; 2) response times to vowel sounds near the category boundary will be longer in the “e-ɛ” task when compared to the “i-e” task.

Methods

Participants

Forty-two (20 females) native speakers of Brazilian Portuguese (spoken in the central region of Minas Gerais) aged between 18 and 34 years participated in this study as volunteers. All of them were right-handed and reported no history of neurological or auditory disorders. Each participant signed a free informed consent form according to the requirements of the Committee on Ethics in Research of the Universidade Federal de Minas Gerais (CAAE: 18350913.2.0000.5149).

Stimuli

Vowel sounds were generated by means of a version of the KLSYN88 synthesizer (Klatt & Klatt, 1990) implemented by the software *Praat* (Boersma & Weenink, 2012; and presented in Weenink, 2009). Twenty-eight vowel sounds were synthesized in such a way as to vary in equal steps along a continuum from /i/ to /e/ to /ɛ/ – based on data from adult male speakers of Brazilian Portuguese (Escudero, Boersma, Rauber, & Bion, 2009; Rauber, 2008). This vowel *continuum* corresponds to a straight-line segment in the space defined by three orthogonal axes representing frequency values of the three lower formants (F_1 , F_2 , and F_3 ; see Figure 2) expressed in the psychoacoustic Bark scale (according to Traunmüller, 1990).

In order to simulate voicing, a glottal flow waveform was generated with duration of 150 ms, with fundamental frequency (F_0) linearly varying from 120 to 90 Hz. These values determining the duration and pitch contour of the stimuli were selected such that the synthesized isolated vowels sounded similar to natural vowels produced by a male speaker (for duration and F_0 measurements from Brazilian Portuguese speakers, see, e.g., Rauber, 2008). A cascade of eight formant filters then modeled vocal tract filtering. The three lower formant frequencies varied along the vowel *continuum* in fixed steps of 0.13 (F_1), 0.06 (F_2) and 0.03 (F_3) Bark, from 2.01 to 5.39 (F_1), 14.20 to 12.52 (F_2), and 15.69 to 14.76 Bark (F_3) – or, in Hertz, from 205 to 555, 2390 to 1860, and 3000 to 2600 Hz. The remaining five formant frequencies

were the same for all sounds (3900, 4900, 5900, 6900 and 7900 Hz). Bandwidths of F_1 , F_2 , and remaining formants were set at 60, 90 and 150 Hz, respectively. Fade-in and fade-out were applied to the first and last 20 ms of each sound. Root mean square (rms) intensity was then scaled to 70 dB (SPL).

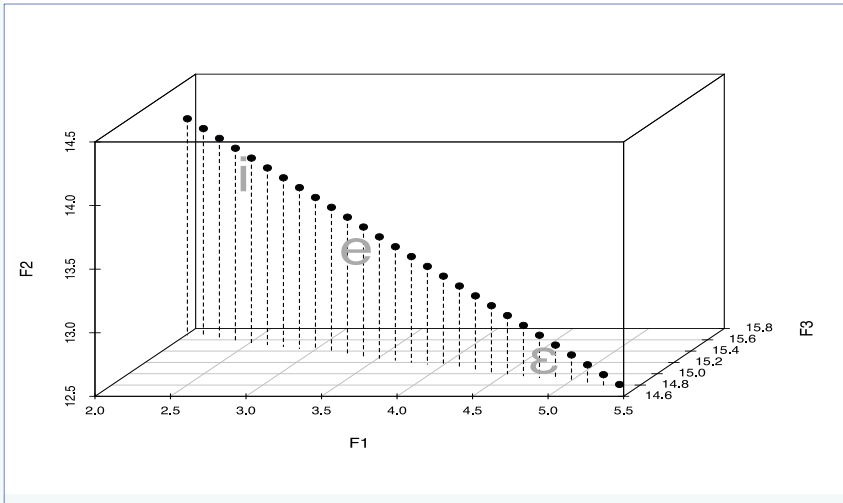


Figure 2 – F_1 , F_2 and F_3 formant frequencies of the vowel sounds used as stimuli in the “i-e” and “e-ε” classification tasks. The sounds varied in equal Bark steps along the [i-ε] continuum.

Experimental design and procedure

Participants performed two classification tasks, each composed of 432 two-alternative forced choice trials in which participants were asked to listen to a sound and indicate, by pressing a button, which of two vowel categories the sound best fits in. Response alternatives correspond to the vowel categories /i/ and /e/ in one of the tasks, and to the categories /e/ and /ε/ in the other. A different subset of sounds was used for each task. The 18 sounds nearest to the /i/ end of the continuum were used as stimuli in the “i-e” classification task, while for the “e-ε” task, the 18 sounds nearest to the /ε/ end were used. Each task comprised 24 repetitions of each sound presented in pseudorandom order (no consecutive trials with the same sound). A 1 s interval was

inserted between the response and the presentation of the next sound. Participants were allowed to pause after every 144 trials.

The order of the two tasks was counterbalanced across participants and each task was preceded by 18 practice trials (one for each sound in the task) presented in random order. Participants were instructed to keep their index fingers on the two response buttons and respond immediately after hearing a sound. Classification responses and response times were recorded. Response times below 180 ms and above 1500 ms were excluded from analysis.

Analysis

Data from two participants (a male and a female) were excluded from the analysis because they did not perform the tasks according to the instructions. In a first step of the analysis, a logistic regression curve was fit to the classification data from each participant and task (“i-e” and “e-ε”). From this curve it is possible to estimate the boundary between the two response categories – defined as the point in the vowel continuum² that corresponds to a response probability $p = 0.5$ (see Figure 3). This boundary point can be found by substituting $p = 0.5$ into the logistic regression equation

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x$$

where p is the probability of one of the response alternatives, x is the independent variable corresponding to the vowel continuum, and β_0 and β_1 are the intercept and slope coefficients, respectively. Therefore, the value of x corresponding to the category boundary is given by $-\beta_0/\beta_1$. Estimations of the boundaries between /i/ and /e/ and between /e/ and /ε/ are thus assigned to each participant.

The focus of the present work is on the sharpness of distinctions between categories, rather than on absolute values such as formant frequencies at the category boundaries. Particularly, we intend to

2. Note that the vowel continuum can be understood as a continuous independent variable that results from a linear combination of F_1 , F_2 , and F_3 .

compare the /i/-/e/ and /e/-/ε/ distinctions in terms of how sharply (and how rapidly) the vowel categories are separated. Thus, in order for results from the “i-e” and “e-ε” tasks to be comparable, it is important that the measures to be compared be made with reference to the category boundaries.

For testing the hypothesis that the /i/-/e/ distinction is sharper than the /e/-/ε/ distinction, d' (Green & Swets, 1974; Macmillan & Creelman, 2005) was used as a measure of the discriminability between the sound at the second level above and the sound at the second level below the category boundary – s_a and s_b , respectively (see Figure 3 for an example). Thus, for the two classification tasks,

$$d'_{i/e} = z(p(r_e|s_a)) - z(p(r_e|s_b))$$

$$d'_{e/\varepsilon} = z(p(r_\varepsilon|s_a)) - z(p(r_\varepsilon|s_b))$$

where $p(r_j|s_k)$ is the proportion of responses r_j to the sound s_k , and $z(p)$ is the inverse of the cumulative normal distribution (proportions of 0.00 and 1.00 were converted to 0.01 and 0.99, respectively). Note that the index j represents the response “e” in the “i-e” task and the response “ε” in the “e-ε” task. Thus, as used here, d' expresses a “classification distance” between two sounds (s_a and s_b), and results from the subtraction between the z-transformed probabilities of assigning them to the same response category. It is used in categorical perception studies as an estimate of discriminability assuming that two stimuli are discriminable to the extent that they are assigned to different categories (Gerrits & Schouten, 2004; Macmillan, Kaplan, & Creelman, 1977; Schouten, Gerrits, & van Hessen, 2003; Silva & Rothe-Neves, 2009). The d' values from the “i-e” and “e-ε” tasks were compared by means of the Wilcoxon signed-rank test.

In order to test the hypothesis that response times are longer for the /e/-/ε/ distinction as compared to the /i/-/e/ distinction, the mean of the response times to the four sounds around the category boundary (from s_a to s_b) was used as the dependent measure. The Wilcoxon signed-rank test was used for the comparison between the two tasks.

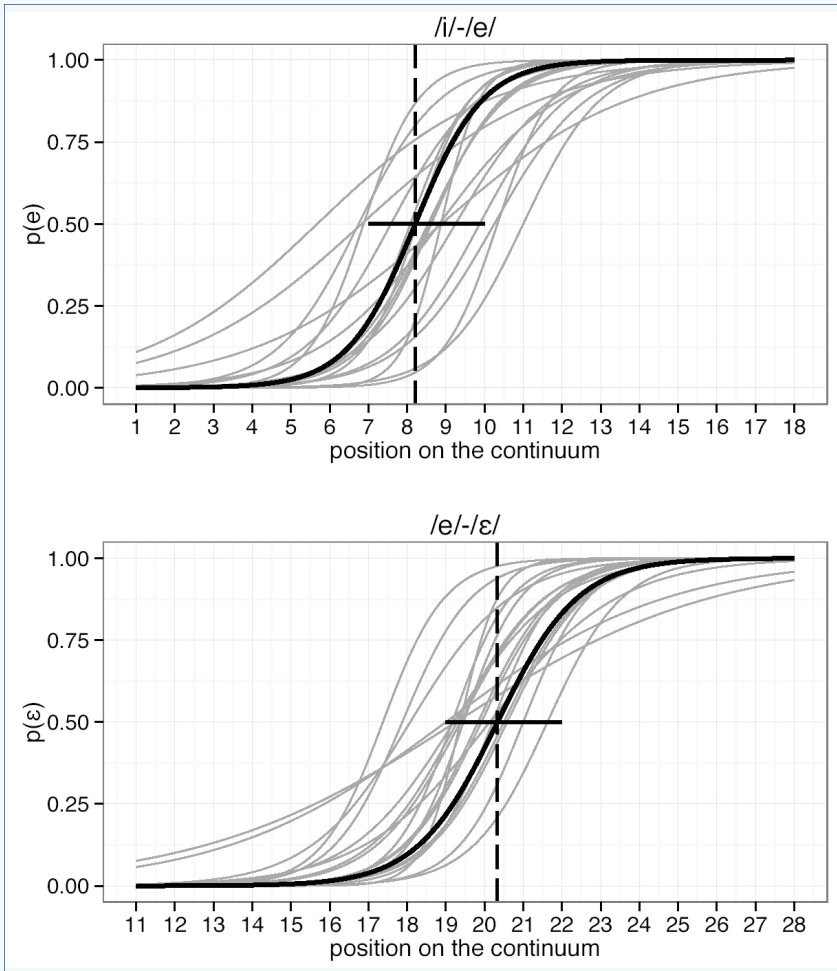


Figure 3 – Fitted logistic regression curves for tasks “i-e” (top) and “e-ε” (bottom). The curve for one of the participants is depicted in black as an example; the dashed vertical line indicates the category boundary for that participant, and the horizontal thick line represents the range of the continuum from which response time and d' measures were obtained (two steps below and two steps above the boundary). Fitted curves for the remaining participants are also shown (thin gray lines).

Results

The fitted logistic regression curves for all participants are shown in Figure 3. For the sake of illustration, the category boundaries and

the s_a - s_b ranges (both defined on a participant-by-participant basis) for a representative participant are also shown³. The mean value of the category boundary estimate fell between positions 8 and 9 for the /i/-/e/ distinction, and between positions 19 and 20 for the /e/-/ɛ/ distinction. The boundary estimates and the corresponding formant values – obtained by linear interpolation in the $F_1 \times F_2 \times F_3$ Bark space – are shown in Table 1. The boundary values varied across participants between 5.68 and 11.00 for the /i/-/e/ distinction (mean 8.28; standard deviation 1.13) and between 17.24 and 21.63 for the /e/-/ɛ/ (mean 19.23; standard deviation 0.99). Participants also varied appreciably regarding the slopes of the regression curves. Odds ratios⁴ from 1.48 to 10.31 were observed in the “i-e” task (80% of the cases, between 1.57 and 4.19; median 2.69) and from 1.26 to 7.83 in the “e-ɛ” task (80% between 1.61 and 3.48; median 2.62).

Table 1 – Category boundary (Mean \pm Standard Error across participants) and the corresponding formant frequencies.

<i>Distinction</i>	<i>Category Boundary</i> (Mean \pm SE)	<i>Formant Values</i>	
<i>/i - e/</i>	8.28 \pm 0.18	F ₁	284 \pm 2.1 Hz
		F ₂	2233 \pm 3.7 Hz
		F ₃	2886 \pm 2.7 Hz
<i>/e - ɛ/</i>	19.23 \pm 0.15	F ₁	430 \pm 2.1 Hz
		F ₂	2017 \pm 2.9 Hz
		F ₃	2723 \pm 2.2 Hz

The comparisons between /i/-/e/ and /e/-/ɛ/ distinctions regarding both d' and response time measures are summarized in Table 2 and Figure 4. Significantly larger d' values (Wilcoxon signed rank test, $V = 536.5$; $p < .05$) were observed in the “i-e” task ($Mdn = 2.88$) than in the “e-ɛ” task ($Mdn = 2.33$). Regarding response-time measures, a

3. Since logistic regression curves were fitted individually for each participant, presenting all such curves and selecting a representative one in order to illustrate the analysis procedure is more appropriate for the current purposes than plotting a regression curve fitted to data averaged across participants.

4. Here, the odds ratio e^{β} represents the ratio by which the odds of selecting a particular response alternative change for a unit change in position along the vowel continuum.

significant difference was also found between the two tasks (Wilcoxon signed rank test, $V = 84$; $p < .001$). Responses were faster for “i-e” ($Mdn = 596$ ms) when compared to “e-ε” ($Mdn = 647$ ms).

Table 2 – Median d' and response times (RT) for distinctions /i-e/ and /e-ε/.

	Distinction		<i>p</i> value	effect size (<i>r</i>)
	/i - e/	/e - ε/		
d' (median)	2.88	2.33	.042	-0.32
RT (median)	596 ms	647 ms	< .001	-0.75

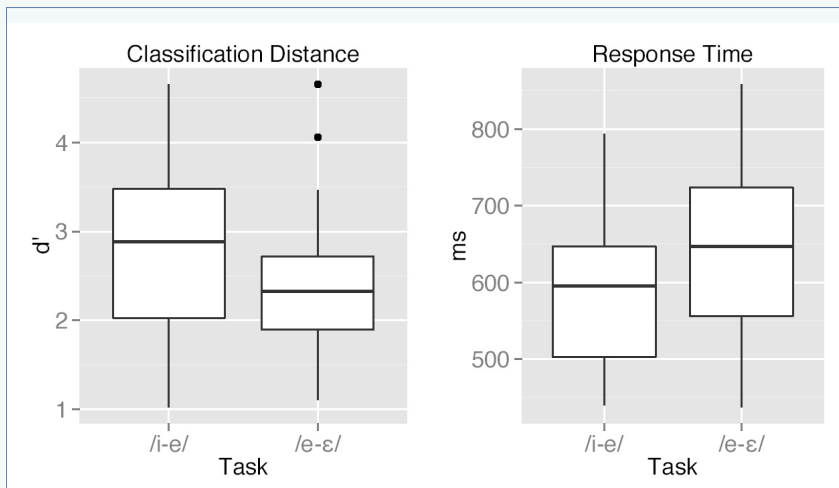


Figure 4 – Boxplots for d' and response times in tasks “i-e” and “e-ε”.

Discussion

Significantly larger d' values and faster responses were observed for the distinction between /i/ and /e/ compared to that between /e/ and /ε/. These results indicate that, at least among front vowels in Brazilian Portuguese, the distinction between mid-high and mid-low vowels is more complex and involves higher perceptual processing costs than the more basic distinction between mid and high vowels. Some considerable variability across participants was observed regarding

both the boundary positions along the continuum and how sharply the two response categories contrasted (as assessed by slopes, odds ratios, and d'). In the latter case, this may reflect performance differences likely associated with limitations of the perceptual system, attention, and motivation. The variability in boundary values reflects individual differences in vowel recognition and/or in decision criteria used by participants to select a response alternative. Importantly, although the participants were Brazilian Portuguese native speakers from a region of the state of Minas Gerais, influences of dialectal factors cannot be ruled out. In any case, regardless of such across-participant variability, nonparametric statistical tests for within participant comparisons revealed significant effects supporting the tested hypotheses on both d' and response time.

As a sensitivity index, d' is more widely used in the analysis of detection and discrimination judgments, but it can also be estimated from the results of two-alternative forced choice classification tasks such as the one employed in the present study (Gerrits & Schouten, 2004; Macmillan et al., 1977; Schouten et al., 2003; Silva & Rothe-Neves, 2009). In such cases, d' provides a measure of “classification distance” that can be understood as reflecting the discriminability between two stimuli in the absence of within-category information, i.e., when the only available information is about the distinction between two categories – analogous to the well known “predicted discrimination” scores used in categorical perception studies to predict discrimination performance from classification responses (Lieberman, Harris, Hoffman, & Griffith, 1957). Here, d' was calculated in such a way as to index classification distance in the region of the boundary between categories (the segment of the vowel continuum between sounds s_a and s_b). As it was larger for the “i-e” task, d' data indicate that the categorization processes involved result in categories that are comparatively less distinct from each other in the case of the opposition between /e/ and /ɛ/. The latter is also associated with longer response times, suggesting that the /e/-/ɛ/ distinction involves processing of more complex information.

These results converge with cross-linguistic typological observations (Diehl, 2008; Schwartz et al., 1997a, 1997b), with some proposals regarding underlying representations of vowels in Brazilian

Portuguese (Lee, 2010; Nevins, 2012), and descriptions of seven-vowel systems such as those of Portuguese (Bisol, 2003; Mateus & d'Andrade, 2000; Wetzels, 2011), Catalan (Wheeler, 2005), and Italian (Krämer, 2009), which show that mid-high:mid-low distinctions are particularly prone to neutralization and variation. Previous results by Silva and Rothe-Neves (2009) on perceptual categorization of back vowels /u o ɔ/ in Brazilian Portuguese are consistent with the present findings in suggesting sharper perceptual distinctions between mid and high than between mid-low and mid-high vowels. Furthermore, they showed that responses in a two-interval forced choice discrimination task were more closely related to classification responses when the stimulus set was composed of sounds from a /u/-/o/ continuum than from a /o/-/ɔ/ continuum. There is therefore some evidence supporting the generalization of the present findings to other vowels. However, response times were not measured in that study (Silva and Rothe-Neves, 2009). On the other hand, the present study did not include discrimination tasks. Hence, response times associated with back vowel distinctions and the relation between classification and discrimination along the /i/-/e/-/ɛ/ continuum should be examined in future studies. Importantly, comparisons between the cases of front and back vowels may shed light on typological observations suggesting that languages tend to prefer front vowel distinctions to back vowel distinctions (Maddieson, 1984; Joanisse & Seidenberg, 1997). Also of interest would be to explore the perception of low:mid and mid-high:mid-low vowel oppositions in different phonological contexts (such as stressed versus pre-stressed syllables), in different languages, and using other measures than the ones provided here – particularly those that can be obtained without any task requirements and in the absence of attention to the stimuli, such as the mismatch negativity brain response (Näätänen, 2001; Silva & Rothe-Neves, 2014).

Conclusion

The present results provide experimental evidence that the distinction between the vowel categories /e/ and /ɛ/ is less sharp and requires more processing effort than that between /i/ and /e/. This is in agreement with the idea that phonological oppositions among

vowels differ in degree of complexity, which would be reflected in typological patterns, phonological representations, speech perception and acquisition. While previous experimental results suggest that the present findings can be extended to the corresponding back vowels /u o ɔ/, further investigations are required to verify whether they would be replicated with back vowels and other speech sound classes.

Recebido em novembro de 2014

Aprovado em setembro de 2015

E-mails: silvadmr27@yahoo.com.br

rothe-neves@ufmg.br

References

- BISOL, Leda. 2003. Neutralização das Átonas. *D.E.L.T.A.* **19(2)**: 267-276.
- BOERSMA, Paul & David WEENINK. 2012. Praat: doing phonetics by computer (Version 5.3.23). Retrieved from <http://www.praat.org>
- BONILHA, Giovana Ferreira Gonçalves. 2004. Sobre a aquisição das vogais. In Regina R. LAMPRECHT. (ed.). *Aquisição fonológica do português: Perfil de desenvolvimento e subsídios para a terapia*. Porto Alegre: ARTMED.
- BOSCH, Laura & Núria SEBASTIÁN-GALLÉS. 2003. Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of Life. *Language and Speech* **46(2-3)**: 217-243.
- DIEHL, Randy L. 2008. Acoustic and auditory phonetics: the adaptive design of speech sound systems. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **363(1493)**: 965-978.
- DRESHER, B. Elan. 2009. *The Contrastive Hierarchy in Phonology*. Cambridge: Cambridge University Press.
- ESCUDERO, Paola; Paul BOERSMA; Andréia Schurt RAUBER & Ricardo A. H. BION. 2009. A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. *The Journal of the Acoustical Society of America* **126(3)**: 1379-1393.
- FIKKERT, Paula. 2005. From Phonetic Categories to Phonological Features Specification: Acquiring the European Portuguese Vowel System. *Lingue e Linguaggio* **2**: 263-280.
- GERRITS, Ellen & M. E. H. SCHOUTEN. 2004. Categorical perception

- depends on the discrimination task. *Perception & Psychophysics* **66(3)**: 363-376.
- GREEN, David M. & John A. SWETS. 1974. *Signal detection theory and psychophysics*. Oxford: Robert E. Krieger.
- HARRIS, John. 1994. *English Sound Structure*. Cambridge: Blackwell.
- _____. & Geoff LINDSEY. 2000. Vowel patterns in mind and sound. In: Noel BURTON-ROBERTS, Philip Carr & Gerard DOCHERTY (eds.). *Phonological knowledge: Conceptual and empirical issues*. New York: Oxford University Press.
- JOANISSE, Marc F. & Mark S. SEIDENBERG. 1997. [i e a u] and sometimes [o]. Perceptual and computational constraints on vowel inventories. In: Proceedings of the NINETEENTH ANNUAL CONFERENCE OF THE COGNITIVE SCIENCE SOCIETY. Vol. 19. Mahwah.
- KLATT, Dennis H. & Laura C. KLATT. 1990. Analysis, synthesis, and perception of voice quality variations among female and male talkers. *The Journal of the Acoustical Society of America* **87(2)**: 820-857.
- KRÄMER, Martin. 2009. *The Phonology of Italian*. Oxford; New York: Oxford University Press.
- KUHL, Patricia. K. 2004. Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience* **5(11)**: 831-843.
- _____.; Barbara T. CONBOY; Sharon COFFEY-CORINA; Denise PADDEN; Maritza RIVERA-GAXIOLA & Tobey NELSON. 2008. Phonetic learning as a pathway to language: new data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **363(1493)**: 979-1000.
- LEE, Sewng-Hwa. 2010. Contraste das vogais no PB e OT. *Estudos Linguísticos* **39(1)**: 35-44.
- LIBERMAN, Alvin M.; Katherine S. HARRIS; Howard S. HOFFMAN & Belver C. GRIFFITH. 1957. The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology* **54(5)**: 358-368.
- MACMILLAN, Neil. A. & C. Douglas CREELMAN. 2005. *Detection theory: A user's guide*. Mahwah: Lawrence Erlbaum Associates.
- _____.; Howard L. KAPLAN & C. Douglas CREELMAN. 1977. The psychophysics of categorical perception. *Psychological Review* **84(5)**: 452-471.
- MADDIESON, Ian. 1984. *Patterns of Sounds*. New York: Cambridge University Press.
- _____. 1999. Phonetic universals. In: William J. HARDCASTLE & John LAVER (eds.). *The handbook of phonetic sciences*. Oxford: Blackwell Publishers.

- MATEUS, Maria Helena & Ernstod' ANDRADE. 2000. *The Phonology of Portuguese*. New York: Oxford University Press.
- MAZUKA, Reiko; Mihoko HASEGAWA & Sho TSUJI. 2014. Development of non-native vowel discrimination: Improvement without exposure. *Developmental Psychobiology* **56(2)**: 192-209.
- NÄÄTÄNEN, Risto. 2001. The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology* **38(1)**: 1-21.
- NEVINS, Andrew. 2012. Vowel lenition and fortition in Brazilian Portuguese. *Letras de Hoje* **47(3)**: 228-233.
- POLKA, Linda & Ocke-Schwen BOHN. 2003. Asymmetries in vowel perception. *Speech Communication* **41(1)**: 221-231.
- _____ & Ocke-Schwen BOHN. 2011. Natural Referent Vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics* **39(4)**: 467-478.
- RAUBER, Andéia Schurt. 2008. An acoustic description of Brazilian Portuguese oral vowels. *Diacrítica, Ciências Da Linguagem* **22(1)**: 229-238.
- RECASENS, Daniel & Aina ESPINOSA. 2009. Dispersion and variability in Catalan five and six peripheral vowel systems. *Speech Communication* **51(3)**: 240-258.
- SCHOUTEN, Bert; Ellen GERRITS & Arjan J. VAN HESSEN. 2003. The end of categorical perception as we know it. *Speech Communication* **41(1)**: 71-80.
- SCHWARTZ, Jean-Luc; Louis-Jean BOË; Nathalie VALLÉE & Christian ABRY. 1997a. Major trends in vowel system inventories. *Journal of Phonetics* **25(3)**: 233-253.
- _____ ; Louis-Jean BOË; Nathalie VALLÉE & Christian ABRY. 1997b. The Dispersion-Focalization Theory of vowel systems. *Journal of Phonetics* **25(3)**: 255-286.
- SILVA, Daniel Márcio R. & Rui ROTHE-NEVES. 2009. An experimental study of the perception of the contrast between the back mid vowels of Brazilian Portuguese. *D.E.L.T.A.* **25(2)**: 319-345.
- _____ & Rui ROTHE-NEVES. 2014. Respostas evocadas de incongruência a categorias na percepção da fala. *Letras de Hoje*, **49(1)**: 66-75.
- TRAUNMÜLLER, Hartmut. 1990. Analytical expressions for the tonotopic sensory scale. *The Journal of the Acoustical Society of America* **88(1)**, 97-100.
- WEENINK, David. 2009. The Klatt Grid speech synthesizer. In *Proceedings of Interspeech 2009: speech and intelligence*. Brighton, UK: International Speech Communication Association.

- WERKER, Janet F. & Richard C. TEES. 2005. Speech perception as a window for understanding plasticity and commitment in language systems of the brain. *Developmental Psychobiology* **46(3)**: 233-251.
- WETZELS, Leo. 2011. The representation of vowel height and vowel height neutralization in Brazilian Portuguese. In: John A. GOLDSMITH; Elizabeth HUME & Leo WETZELS (eds.). *Tones and Features: Phonetic and Phonological Perspectives*. Berlin: Walter de Gruyter.
- WHEELER, Max W. 2005. *The Phonology of Catalan*. New York: Oxford University Press.

