CHILDREN'S PERCEPTUAL AUDITORY PERFORMANCE IN IDENTIFYING PHONEMIC CONTRASTS

Larissa Cristina BERTI*

▪ ABSTRACT: This study aimed at investigating children's perceptual auditory performance in identifying phonemic contrast in Brazilian Portuguese (henceforth BP). The hypothesis is that the perceptual auditory acquisition develops in a gradual fashion, following a systematic acquisition order. We performed four identification tasks using the instrument PerceFAL with 66 children (of both genders) between 4-5 years old. The task relied on the presentation of an acoustic stimulus, through earphones, and the choice of an image corresponding to the word shown, having two image possibilities available on the computer screen. We compared both the stimulus length of time and reaction time of children automatically through the aid of the software PERCEVAL. The children's perceptual auditory performance occurred gradually and depended on the phonemic class. A greater accuracy regarding the phonemic contrast identification seems to follow the sequence: vowels, sonorants, stops e fricatives. The reaction time for the correct answers was shorter than that of the incorrect answers (except for the vowel class). From the perceptual maps, we verified that, within the vowel class, the anterior-posterior parameter plays an important role in perceptual salience. For the obstruents and sonorants (nasal and liquid), the acoustic cues that characterize voicing (in the case of obstruents) and the articulation mode (in the case of sonorants) are perceptually more robust than the cues from the point of articulation. Although speech perception should not be reduced to a mere sensory interpretation, the acoustic cues of speech segments exert influence on their categorization.

▪ KEYWORDS: Speech perception. Language Acquisition. Acoustic Phonetic. Brazilian Portuguese.

Introduction

Throughout the phonetic and phonological acquisition of a language, researchers usually highlight the role performed by children´s articulatory and auditory skills, besides the sensory-motor connections that underlie such a process (MUNSON et al., 2005; GATHERCOLE, 2006; HARDCASTLE et al., 2010; PANNETON; NEWMAN, 2011; as the latest study references).

* UNESP – State São Paulo University. Faculty of Philosophy and Science. Speech Language Pathololy and Audiology Department. Marília – SP – Brazil. 17525-900 – berti.larissa@gmail.com

In the Brazilian context, studies developed on phonemic acquisition have primarily tackled speech production investigation, that is, when and how children reach the target production in their language (LAMPRECHT et al., 2004). There are not, as far as our preliminary bibliographic review indicated, studies that focus on the process of phonological contrasts acquisition taking perception into account.

In the international literature, there have been reports of researches on children´s phonemic contrasts domain, especially regarding the skills of discriminating and identifying such contrasts since the late 1940s, with a great surge in the 1970s and 1980s.

We understand, according to Hazan and Barrett (2000), that the discrimination task assesses the skill to perceive difference between two sounds, and the listener does not have to categorize the sounds in question – the comparison is performed *in praesentia*.

Concerning children´s auditory perceptual development, previous studies have shown that 1-month old babies are able to discriminate phonemic contrasts of their native language, and also phonemic contrasts of other languages (EIMAS et al., 1971; STREETER, 1976; TREHUB, 1976; EILERS; GAVIN; OLLER, 1982; MEHLER, 1985, for a more detailed review).

Likewise, previous studies in literature have shown that children´s skills at discriminating non-native contrasts rapidly diminish in childhood due to linguistic experience. At the same time, there is an increase in accuracy to discriminate phonemic contrasts in their language (WERKER; POLKA, 1993, for a review of studies that establish such an observation). In other words, the ability related to children´s phonemic discrimination is gradually impaired in the following pattern: from discrimination of potential phonemic contrasts (non-native and native) to discrimination of contrasts in their native language.

In a series of investigations, Werker and colleagues (WERKER et al., 1981; WERKER; TEES, 1984a, b; WERKER; LALONDE, 1988; BEST; McROBERTS, 1989; BEST, 1994) have demonstrated that the decline in the ability to discriminate non-native contrasts takes place along the child´s first year.

What is particularly fascinating in the described findings relies on the fact that the decline in the perceptual ability does not seem to occur in the same fashion for all non-native contrasts. The experimental results indicate that the failure to distinguish certain contrasts occurs before the failure related to other contrasts – which allows to hypothesize that the decline of auditory perceptual development is gradual and occurs in a systematic order.

As possible explanations for the impairment of non-native auditory perceptual discrimination, Best and McRoberts (1989), Best (1993), Best (1994), Werker and Tees (2002), have suggested that such a decline might reflect the first stage of children´s phonological acquisition. However, the authors do not discuss which developing phonological aspect could be responsible for the change in the auditory perceptual domain.

In order to acquire the phonology of a language, from an auditory perceptual perspective, children have to learn and discriminate not only the sound patterns in their language, but also organize them effectively into appropriate phonemic categories (HAZAN; BARRETT, 2000). The latter ability is known in the speech perception literature as "phoneme categorization" or "phoneme identification", in which the listener is required to categorize sounds, that is, the comparison is performed *in absentia*[1].

Regarding the development of the identification ability, researchers have demonstrated that children´s ability to identify (or categorize) phonological contrasts of their native language develops not only gradually but systematically in terms of acquisition (SHVACHKIN, 1948; GARNICA, 1973; EDWARDS, 1974; BARTON, 1980; BROWN; MATTHEWS, 1993, 1997).

In a classic study, for example, Shvachkin (1948) investigated the development of Russian-speaking children to identify certain contrasts and found out they tend to be better when compared to others. Thus, there might be an auditory perceptual acquisition order, similar to that described by Jakobson (1968)[2] when it comes to speech. According to him, children undergo two great periods in their auditory perceptual development: in the first place, a distinction among vowels (discrimination and identification) and in the second place, among consonants. In the second period, in turn, there might be 11 distinct stages, namely: (1) distinction between presence x absence of consonants (ex: /ok/ vs. /bok/); (2) distinction between sonorants and what the author called articulated obstruents (ex: /m/ vs. /b/); (3) distinction between palatalized consonants x non-palatalized consonants; (4) distinction between sonorants (nasal vs. liquid); (5) distinction between sonorants and what the author called non-articulated obstruents (ex: /l/ vs. /x/); (6) distinction between labials and coronals; (7) distinction between stops and fricatives; (8) distinction between coronals and dorsals; (9) distinction between voiced and voiceless; (10) distinction between strident sibilants and non-strident sibilants and, finally, (11) distinction between liquids.

Further studies (EDWARDS, 1974; BARTON, 1980; BROWN; MATTHEWS, 1993; HAZAN; BARRETT, 2000; PATER et al., 2004) have systematically reinforced three broad trends for auditory perceptual phonemic acquisition, namely: (a) seven-year old children are still to finish the process of phonemic contrasts perception; (b) phonemic perception gradually develops, usually with production advancement; (c) the order of perceptual acquisition tends to become uniform among the world languages, but it is not universal. Trends in differentiations are common in varied languages (e.g. the distinction between consonants vs. vowels, oral vs. nasal). Nevertheless, there is

---

1 In order to learn in depth the concepts of auditory discrimination and identification, as well as the tasks that assess skills, we recommend the following reading: Gerrits (2001).

2 Jakobson (1968) establishes that phonological development, in terms of production, derives from an original situation and tends to become differentiation and separation. The first great opposition occurs with consonants and vowels, followed by the oral vs. nasal opposition, for consonants and vowels. Afterwards, there is a distinction between labials and dentals for consonants and the distinction between broad and narrow vowels, which anticipates the opposition between front and back consonants. The presence of fricatives suggests the presence of stops. Likewise, the presence of affricates suggests the presence of fricatives. Round vowels appear after non-round vowels. Back consonants appear after the presence of front consonants, as voiced consonants occur after voiceless consonants. Liquid consonants are the last to appear and the distinction between lateral and non-lateral liquids are seen to be acquired later in the languages that harbor them.

divergence in opinion as for the appearance of distinction between labials vs. dentals vs. dorsals or even between stops vs. fricatives vs. affricates.

Since the abilities to discriminate and identify play a fundamental role in the phonemic acquisition process, it is relevant to understand how they modify and evolve throughout the process.

Hence, this study aims at investigating children´s auditory perceptual performance in the identification of phonemic contrasts in the BP. More specifically, this study will verify the level of difficulty a given phonemic class poses; and, within each class, an indication of similarity/dissimilarity among the phonemes, with a proposal of perceptual maps for each phonemic class.

Considering existing studies in international literature, our hypothesis relies on the idea that auditory perceptual acquisition gradually evolves, so as to follow a systematic acquisition order.

The scientific gains for both Linguistics and Speech Therapy are the following: (a) contribution for the understanding of acquisition and development of children´s auditory perception with typical language development; (2) generation of data of auditory perception phonemic contrasts in the BP; (3) contribution for the study of production and speech perception.

Method

Participants

This study was approved by the Research Ethics Committee at the Universidade Estadual Paulista (UNESP/Marília), having received the following number 132/2010.

Data from 140 children aged between 4 and 8 years old were collected. The criteria to include children in the sample selection were: typical language development and absence of otological and/or auditory conditions, which was confirmed by a previous auditory triage; whereas the criterion to exclude was the participation of each child in the four identification experiments.

In the end, the sample contained 66 children, of both genders, aged between 4 and 5 years old. The children were selected from a municipal school in the city of Marília, in the state of São Paulo. The children´s parents and tutors signed an Informed Consent and allowed the children to participate in the research.

Material

We used a phonemic contrasts identification instrument – PerceFAL (BERTI, 2011) and the software PERCEVAL (Perception Evaluation Auditive & Visuelle) (ANDRÉ et al., 2009).

PerceFAL is composed of a subset of four experiments: (a) PerceVog (which evaluates the identification of stressed vowels); (b) PerceOcl (which evaluates the identification among stops); (c) PerceFric (which evaluates the identification among fricatives); (d) PerceSon (which evaluates the identification among sonorants).

This instrument deals with two-syllable words with stress on the first syllable, familiar to the children, containing the 19 consonantal phonemes of BP.

The words were chosen according to the following criteria: (1) contrast the six BP stops, so as to compose minimum pairs of words; (2) be represented by means of pictures; (3) belong to children vocabulary; (4) belong to a word list from a previous study (MOTA, 2001).

With PerceVog, it is possible to evaluate the identification of stressed vowels from a set of 42 contrasting pairs (through combinatorial analysis: 7 stressed vowels vs. 6 other distinctive vowels); with PerceOcl, it is possible to evaluate the identification of stops, from 30 minimal pairs (6 stops x 5); with PerceFric, it is possible to evaluate the identification of fricatives, considering 30 contrasting pairs (6 fricatives x 5) and; finally, with PerceSon, it is possible to evaluate the identification of the sonorants, from 42 contrasting pairs (7 sonorants (3 nasals and 4 liquids) x 6).

The following Figures (1-4) show the contrasting pairs selected to compose our experiment.

Figure 1 – Minimal word pairs involving the stressed vowels in the PerceVog

Vowel Contrasts	Minimal Pairs
/i/ - /e/	bico–beco
/i/ - /ɛ/	vila–vela
/i/ - /a/	pipa–papa
/i/ - /ɔ/	chique–choque
/i/ - /o/	figo–fogo
/i/ - /u/	lixo–luxo
/e/ - /ɛ/	feira–fera
/e/ - /a/	pera–para
/e/ - /ɔ/	feira–fora
/e/ - /o/	seco–soco
/e/ - /u/	seco–suco
/ɛ/ - /a/	berro–barro
/ɛ/ - /ɔ/	cheque–choque
/ɛ/ - /o/	beca–boca
/ɛ/ - /u/	fera–fura
/a/ - /ɔ/	bala–bola
/a/ - /o/	saco–soco
/a/ - /u/	lava–luva
/ɔ/ - /o/	toca–touca
/ɔ/ - /u/	coca–cuca
/o/ - /u/	soco–suco

Source: author’s elaboration.

Figure 2 – Minimal word pairs involving the stops in the PerceOcl

Stop Contrasts	Minimal Pairs
/b/ x /t/	berço-terço
/b/ x /k/	bola-cola
/g/ x /b/	gola-bola
/b/ x /p/	bote-pote
/b/ x /d/	bucha-ducha
/d/ x /g/	danço-ganso
/g/ x /t/	guerra-terra
/p/ x /g/	pato-gato
/p/ x /d/	pente-dente
/p/ x /k/	porta-corta
/t/ x /d/	tia-dia
/t/ x /p/	torta-porta
/k/ x /g/	cola-gola
/k/ x /t/	couro-touro
/d/ x /f/	fada-faca

Source: author's elaboration.

Figure 3 – Minimal word pairs involving the fricatives in the PerceFric

Fricative Contrasts	Minimal Pairs
/f/-/v/	faca-vaca
/f/-/s/	fanta-santa
/f/-/z/	forro- zorro
/f/-/ʃ/	fora-chora
/f/-/ʒ/	faca-jaca
/v/-/s/	vela-sela
/v/-/z/	cavar-casar
/v/-/ʃ/	veia-cheia
/v/-/ʒ/	vaca-jaca
/s/-/z/	caçar-casar
/s/-/ʃ/	sapa-chapa
/s/-/ʒ/	selo-gelo
/z/-/ʃ/	rosa-rocha
/z/-/ʒ/	zangada-jangada
/ʃ/-/ʒ/	xis-giz

Source: author's elaboration.

Figure 4 – Minimal word pairs involving the sonorants in the PerceSon

Sonorant Contrasts	Minimal Pairs
/m/-/n/	mata-nata
/m/-/ɲ/	uma-unha
/m/-/l/	mata-lata
/m/-/ʎ/	comer-colher
/m/-/r/	fumo-furo
/m/-/R/	mata-rata
/n/-/ɲ/	sono-sonho
/n/-/l/	nata-lata
/n/-/ʎ/	fina-filha
/n/-/r/	caneta-careta
/n/-/R/	nata-rata
/ɲ/-/l/	punho-pulo
/ɲ/-/ʎ/	pinha-pilha
/ɲ/-/r/	sonho-soro
/ɲ/-/R/	unha-urra
/l/-/ʎ/	vela-velha
/l/-/r/	pulo-puro
/l/-/R/	lata-rata
/ʎ/-/r/	alho-aro
/ʎ/-/R/	colher-correr
/r/-/R/	caro-carro

Source: author's elaboration.

In order to carry out the identification test, PerceFAL features auditory stimuli, corresponding to audio files of all the words by a typical adult speaker; visual stimuli, graphic pictures of each word; besides scripts[3] to run the identification experiment in the software PERCEVAL.

Experimental procedure

The perceptual experimental procedure consisted of an identification test (also known as forced choice task), composed of three distinct steps: recognition of the experiment words, training and testing.

[3] The scripts of the identification experiments can be sent by the author upon request: berti.larissa@gmail.com.

The recognition step consists of the presentation of visual and auditory input to children so as to verify whether they know the words and/or pictures used in the experiment. Some words in the instrument have easy corresponding pictures (*pato*, *gato*, etc.), whereas some might raise doubts (as in *unha*, *gola*, etc.). After children were familiar with the experiment input, we checked whether they knew the words. A threshold of 80% of correct answers would lead the children to the training step and the perceptual step.

The training step is carried out automatically by the software and aims at enabling the participants to understand the task. This step consists of the perceptual identification task, but the results are not computed. The experiment stimuli are randomized and 10 presentations are selected. Afterwards, the testing step is initiated.

For the identification task, the children were comfortably placed in front of a computer screen (with the software PERCEVAL installed) and used KOSS headphones, inside an acoustic booth. Each child heard (with binaural presentation) one of the minimal pairs word, and had to decide and indicate the corresponding picture. Two pictures were displayed on the computer screen. For example: they heard the word *pote* and two pictures appeared on the computer screen, representing the words *pote* and *bote*. The participant had to decide and indicate the picture that corresponded to the auditory stimulus. Both the time the auditory and visual stimulus was shown and the reaction time were controlled and measured automatically by the software PERCEVAL.

The overall duration of each experiment was approximately 15 minutes per child. Even though the experiments were performed in different days to prevent children from becoming exhausted, the conclusion did not last more than a week.

Analysis criteria

The following analysis criteria were used: a) auditory perceptual accuracy; b) reaction time for right and wrong answers; c) pattern recognition with the aid of multidimensional scaling and a similarity matrix.

Results

Auditory perceptual accuracy

Since one of the goals of the proposed analysis was to detect auditory perceptual accuracy in identifying phonemic contrasts in BP, instead of analyzing only the percentage of correct answers, the data were transformed by using a measure of sensitivity known as d’prime (MACMILLAN; CREELMAN, 1991). This sensitivity index takes into consideration the variation of answers from subjects by adjusting the number of hits (right answers, that is, stimulus A is chosen when the stimulus shown was A) by the number of false alarms (wrong answer, stimulus A is chosen when the stimulus shown was B). D'prime is calculated by the conversion of the proportions of hits (H) and false alarms (FA) into z-score and, afterwards, by the subtraction of the proportions (d'= z(H) –z(FA)).

Perfect accuracy (only hits and no false alarms) would show an infinite d'prime value. Then, the hit and false alarm values were adjusted to a proportion of H=0,99 and FA=0,01, with a nearly perfect accuracy value, d'=4,65. Thus, the closer to 4,65 the d'prime is, the higher the auditory perceptual accuracy.

Table 1 shows d'prime values for phonemic classes.

Table 1 – Auditory perceptual accuracy for phonemic classes

Phonemic Class	Number of Answers (number of contrasting pairs x 66 children)	% of correct answers	d'prime
Vowels	2,772	88.34 (2449/2772)	1.73
Sonorants	2,772	87.01 (2412/2772)	1.56
Stops	1,980	84.04 (1664/1980)	1.43
Fricatives	1,980	75 (1485/1980)	0.93

Source: author's elaboration.

The children's auditory perceptual accuracy was seen to be dependent on the phonemic class, with the following pattern: vowels>sonorants>stops>fricatives. The accuracy values (d'prime) varied from 1.73 to 0.93.

Reaction time for wrong and correct answers

Table 2 shows the comparison between the mean response time for the answers in the phonemic classes.

Table 2 – Comparison between the mean response time for the answers

Phonemic Class	Mean reaction time for right answers (ms)	Mean reaction time for wrong answers (ms)	T-test for independent samples
Vowels	2,158.31 (±221.49)	2,243.83 (±455.80)	t=-1.34, p= 0.17
Sonorants	2,171.17 (±251.99)	2,388.74 (±478.09)	t=-2.92, **p= 0.00**
Stops	2,037.04 (±218.90)	2,200.90 (±543.34)	t=-6.45, **p= 0.00**
Fricatives	2,346.71 (±236.02)	2,411.51 (±346.39)	t=-2.10, **p= 0.03**

Source: author's elaboration.

The reaction time refers to the time elapsed until the children made the decision in the identification task. Since the identification tasks were performed within four phonemic classes (vowels, sonorants, stops and fricatives), the reaction times for wrong and right answers were compared considering each class individually and not the comparison among classes.

According to Table 2, the reaction time mean value for wrong answers was higher than the reaction time mean value for right answers in all phonemic classes. A t-test was run so as to check whether the reaction time mean value for right answers was significantly different from the reaction time mean value for wrong answers. The alpha value established was 0.05. That means that when the p-value is lower than 0.05, the null hypothesis is rejected (there is no difference between the mean reaction times for wrong and right answers) and the alternative hypothesis is accepted (there is difference between the mean reaction times for wrong and right answers). Hence, the inferential statistical analysis showed that mean reaction time for right answers was lower, except for the vowel class.

The final procedure was the verification of the class, similarities and dissimilarities among the phonemes, transforming the distance measures, as proposed by Johnson (1991), so as to create perceptual maps.

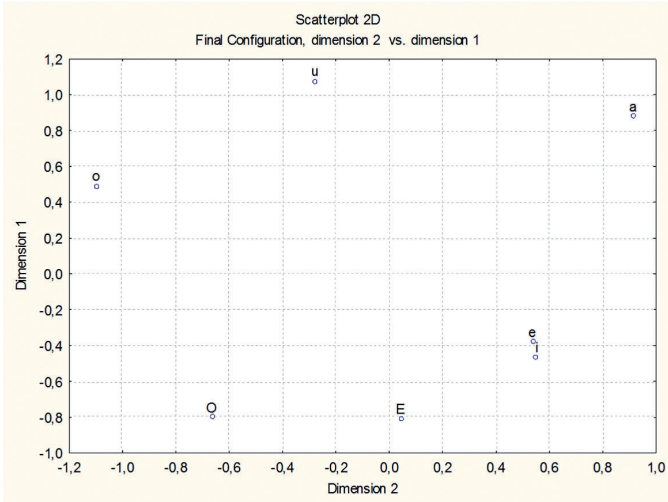
A confusion matrix[4] of each phonemic class was generated by the children's answer pattern and a statistical exploratory analysis was performed (multidimensional scaling and similarity matrix) to verify the similarities/dissimilarities in the contrasting pairs investigated.

Taking into consideration the similarity values for all the investigated contrasting pairs, distance measures were calculated for the pairs, resulting in conceptual maps. For example, in the fricatives class, identification errors involving the pair /s/ vs. /ʃ/ were more frequent than the errors involving the pair /s/ vs. /v/. In other words, the frequency which a sound is confused with another one is caused by similarity. Then, the voiceless coronal fricatives proved to be more similar, for children, than the pair /s/ vs. /v/. The distance measure obtained from the contrasting pairs similarity shows that the distance between voiceless coronal fricatives is lower than the distance between /s/ vs. /v/.

The four conceptual maps are displayed with the phonemic class investigated: vowels, stops, fricatives and sonorants.

[4] A confusion matrix is a notational system used to quantitatively and qualitatively catalog children's answer pattern. Children's wrong and right answers are registered. This type of register provides information regarding lower and higher similarity contrasts in the identification task, as well as the most frequent patterns.

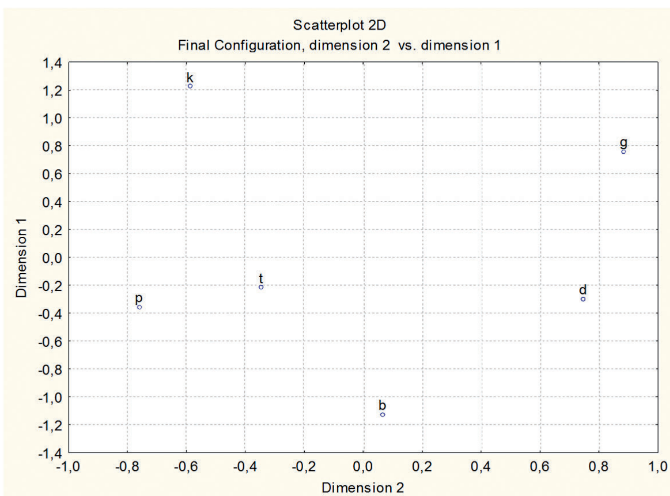
Image 1 – Perceptual Map for the stressed vowels of children



Source: author's elaboration.

The perceptual map shows that there is an unequal distribution among the vowel phonemes in the vowel class, with a pronounced division in the vowel elements considering backness. The front vowels are clustered in the lower right, the back vowels are displayed on the left of the map and the central vowel /a/ is seen in the upper right. Furthermore, within the front vowels, the vowels /e/ and /i/ show a short distance, which reflects high perceptual similarity in children.

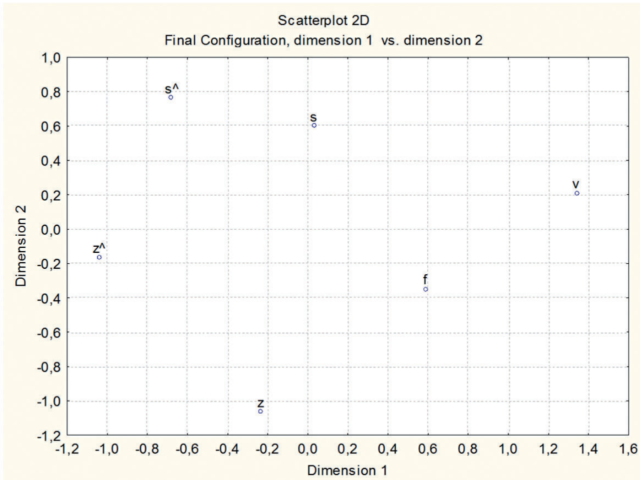
Image 2 – Perceptual map of the stops in children



Source: author's elaboration.

The perceptual map of the stops demonstrates that there is an evident separation between voiceless stops (on the left) and voiced stops (on the right). Additionally, the labial and dental stops are closer when compared to dental and velar stops.

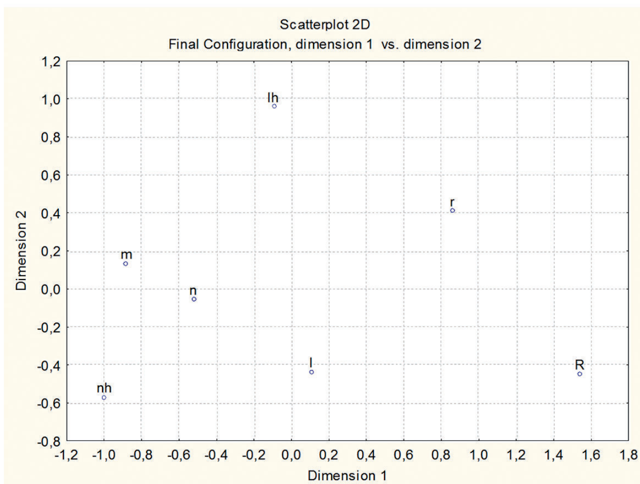
Image 3 – Fricatives perceptual map of children



Source: author’s elaboration.

The fricatives perceptual map shows that in the stops, there is a separation between voiced fricatives (more distributed in the edges) and the voiceless fricatives (central part of the map). The coronal fricatives show a shorter distance among themselves.

Image 4 – Sonorants perceptual map of children



Source: author’s elaboration.

Finally, the observation of the sonorant perceptual map allows to see a distinction between nasal and liquid class. The nasals are clustered in the lower left whereas the liquids are distributed on the right. Within each subset, the nasals are seen to be closer among themselves when compared to the liquids.

Discussion

Auditory perceptual accuracy

Three aspects of the obtained results can be highlighted when it comes to auditory perceptual accuracy. The first refers to the accuracy value (d'prime) shown by the children varying from 0.93 to 1.73. The children in the age range (4-5 years) studied still do not show effective mastery at the contrast identification – all the values were lower than 4.65 (reference value for perfect accuracy).

The fact that four and five-year olds do not show effective mastery in the phonemic identification is consistent with the findings described by Edwards (1974) and Werker and Lalonde (1988), which state that seven-year olds are still to finish the process of phonemic contrasts perception, which suggests that the conclusion and steady perception of phonemic contrasts take place years later. It is feasible to think about not only as a phonological acquisition process in line with phonemic class, but also as an ongoing process throughout development.

The second aspect concerns a gradual domain of the phonemic classes, that is, from a perceptual standpoint, the children's development follows a pattern: vowels>sonorants>stops>fricatives.

The obtained results corroborate those found in international classic studies (SHVACHKIN, 1973; EDWARDS, 1974; BROWN, 1997; 2000; PATER; STAGER; WERKER, 2004), which predict not only a gradual auditory perceptual acquisition, but also a an acquisition hierarchy.

Particularly, the vowels are the first to be differentiated, followed by the consonants. Within the consonant class, the distinction between sonorants x stops comes before a distinction between stops x fricatives, as previously described by Shvachkin (1948).

The phonemic classes with more accuracy in the identification task, vowels and consonants, show a well-defined formation pattern in the acoustic structure (KENT; READ, 1992), characterized by a reinforcement of 300-3000 Hz frequencies. This frequency band features the frequencies perceived by the human ear, and this frequency set is amplified (around 10-12 dB) thanks to physical and physiological characteristics of the ear (LOPES; CAMPOS, 1994; JOHNSON, 1991).

On the other hand, the lower accuracy for the affricate class also reinforces the important interaction between acoustic features and anatomical-physiological features of the human ear for the phonemic identification. From an acoustic perspective, the fricatives are usually present with aperiodic energy distributed in the frequency spectrum according to the length of front cavity resulting from production. More specifically, the shorter the front cavity length in the fricatives, the higher the resonance frequencies (KENT; READ, 1992). In terms of human ear sensibility, frequencies higher than 5,000 Hz do not receive any increment, and are less salient to lower frequencies (JOHNSON, 1997).

All these findings entail the importance of acoustic features of speech sounds for the classification.

Finally, the third aspect refers to the absence of correspondence for the established order in terms of oral acquisition. The liquid class is the last to be acquired in terms of production (JAKOBSON, 1968; LAMPRECHT et al., 2004). The data indicate that the liquids were treated in conjunction with nasals, within the sonorant class, liquids showed a lower percentage of wrong answers (18.57% - 60/323) than the nasals (36.22% - 117/323). Despite the relation between production and speech perception, it is not linear.

In line with Casserly and Pisoni (2010), since speech production cannot be reduced to a mere motor activity, speech perception cannot be reduced to mere sensory interpretation either.

Reaction time for wrong and right answers

Taking into consideration each class, the reaction time for right answers was seen to be inferior to the reaction time for wrong answers, except for the vowels.

Pisoni and Tash (1974) predicted the relationship between perceptual similarity and reaction times in perceptual experiments. According to them, the higher the acoustic difference between two pairs, the faster the subject should answer (the lower the reaction time). On the other hand, the lower the acoustic difference, the higher the time for the subjects to take a decision (the higher the reaction time).

Understanding that the reaction time should be higher for similar segments, we speculate that the perceptual auditory mistakes made by children are likely to involve more similar segments, which would demand a higher reaction time.

It is feasible to understand that the consonantal phonemic contrasts that are linked to more mistakes in the identification task do pose more perceptual auditory similarity, which eventually entails more time for decision taking, in terms of psycholinguistic processing.

The exception for the vowels is justified by their own acoustic features, that is, vowels are segments with more duration and acoustic energy. They also contribute to the reinforcement of frequencies (formants) in a special band for the human ear, facilitating perception (KENT; READ, 1992; JOHNSON, 1991). Then, similarity in the vowel class was seen to be lower than in the consonant class.

Another explanation allows to state that the higher reaction time for wrong answers might stem to the absence of causality in the answers, providing reliance. If the mistakes made by the children in the identification task had happened at random, that is, if the children had guessed the answer, a lower reaction time would have been found for the wrong answers.

Perceptual maps

The children´s perceptual map for the stressed vowels shows an unequal distribution for the vowel phonemes, which reflects different levels of perceptual similarity of the elements. Furthermore, the front vowels /e/ and /i/ were seen to show a shorter distance among them, demonstrating an enormous perceptual similarity by children.

Different levels of perceptual auditory similarity among vowels have been reported in previous studies (POLKA; WERKER, 1994; POLKA; BOHN, 1996; BERTI; ROQUE, 2013). The explanation given by Polka and Bohn (1996) is that the stimuli, in a perceptual domain, are not equally salient. They state that the peripheral vowels serve as a “anchoring place” in the perceptual auditory task, labeled by the authors as natural perceptual magnets.

The data obtained in the children´s perceptual performance regarding the vowel class are in accordance with the Polka and Bohn (1996): it is possible to observe the presence of extreme vowels distributed peripherically throughout the map; The Brazilian Portuguese also seems to have a phonetic similarity for the vowels. The vowel raising involving the mid vowels (TENANI; SILVEIRA, 2008) shows that a language fact is present in perceptual similarity/dissimilarity in children aged 4-5.

Regarding the results of the children´s perceptual map for the stops and fricatives, two items should be pointed out. The first is related to voiced and voiceless obstruents, more evident in the case of stops (with a clear separation in the map) and subtler in the case of fricatives (the voiced sounds are distributed at the ends and the voiceless sounds are in the center of the map).

The second aspect encompasses the role of phonetic cues and the place of articulation of obstruents in the children´s perceptual performance. In the case of the stops, there is a shorter distance between labial and dental stops when compared to labial vs velar and dental vs velar, especially in the voiceless stops. As for the fricatives, the coronal fricatives show a shorter distance.

We hypothesize that phonetic features concerning the place of articulation and voiced obstruents might play a key role in children´s perceptual performance.

The perceptual distinction between voiced and voiceless obstruents has also been described in a previous study (MILLER; NICELY, 1955), which found that voicing is the most salient acoustic cue for English-speaking adults. The acoustic cues that mark voicing are more salient than the cues that mark the place of articulation.

Ferreira-Silva and Pacheco (2011), in a study on the fricative contrast perception, also highlighted the importance of voicing so as to distinguish fricatives.

Concerning the role of the acoustic cues that mark the place of articulation, despite the perceptual differences related to the obstruents place of articulation, it is not fully understood how such differences are perceptually established in adults (MILLER; NICELY, 1955; WANG; BILGER, 1973; WINTERS, 2000).

According to Miller and Nicely (1955), in the places of articulation of the obstruent consonants, the coronals show more auditory perceptual salience. Nevertheless, there are no substantial differences between labial and velar obstruents in terms of auditory perceptual salience. Differently, Wang and Bilger (1973) found that labial and coronal obstruents show, high auditory perceptual salience, whereas velar obstruentes show less salience. More recently, Winters (2000) found that labial obstruents are typically more salient as place of articulation, whereas velar obstruents are described as less salient consonants. Berti et al. (2012) described that labial and coronal stops showed more similarity among children when compared to velar stops.

Children´s perceptual results for the obstruents are similar to those reported by Wang and Bilger (1973). The labials (in the case of the stops) and coronals were seen to have a shorter distance, which suggests more perceptual auditory similarity.

Phonetic features of the obstruents are likely to play a fundamental role in children´s perception.

As for the children´s perceptual map of the sonorants, there is a difference between nasals and liquids. Within each subclass, the nasals are closer than the liquids.

A possible interpretation lies on the phonetic features of the segments in children´s perception. The distinction between nasals and liquids leads to more perceptual similarity among same subclass elements. Acoustically, the nasals are characterized by the presence of a well-defined nasal formant (KENT; READ, 1992), which may explain both the separation of classes and the nasal phonemic proximity.

The findings also corroborate the predictions of Borden et al. (1994), regarding the distinct perceptual saliences of acoustic cues – interpreted here as a type of perceptual hierarchy. The acoustic cues that mark the manner of articulation seem to be more salient than the acoustic cues that mark the place of articulation, as there is a more distinct division between manners (nasal and liquid) than point of articulation.

These findings confirm the hypothesis that there is a perceptual auditory acquisition in the Brazilian Portuguese. Such acquisition occurs gradually and obeys a systematic order in which the idea of phonemic class plays a major role. In the perceptual auditory acquisition, different phonemic classes should be taken into account and within each class, there is an internal hierarchy, where the cues define voicing. The cues to distinguish manner are more salient than the cues that define articulation.

We are aware that our study has an important limitation: the instrument that was used relied on pictures that do not show the same representability and the words were not controlled in terms of frequency in the language, part of speech, vowel context, etc. The children´s perceptual performance might have been influenced.

Conclusions

Children´s perceptual auditory performance occurs gradually and is dependent on phonemic class. The identification of phonemic contrasts seems to follow a pattern: vowels, sonorants, stops and fricatives.

The perceptual maps show that, within the vowel class, backness can play an important role in perceptual salience. For the obstruent class (fricatives and stops) and sonorants (nasals and liquids), the acoustic cues that define voicing (in the case of obstruents) and the manner of articulation (in the case of sonorants) are more perceptually salient than the cues of the place of articulation.

Finally, even though there is a relationship between speech production and perception, it is not viable to assure that such relationship is linear or mirrored. Theories and experimental evidences need to converge so as to find out how production and perception interact in the complex communication act.

The children's perceptual auditory performance in identifying phonemic contrasts will be extended and will consider not only the age range studied, but also the development of longitudinal studies.

BERTI, L. Desempenho perceptivo-auditivo de crianças na identificação de contrastes fônicos. **Alfa**, São Paulo, v.61, n.1, p.79-99, 2017.

▪ *RESUMO: O objetivo deste estudo foi investigar o desempenho perceptivo-auditivo de crianças na tarefa de identificação de contrastes fônicos do Português Brasileiro (doravante PB). A hipótese foi a de que a aquisição perceptivo-auditiva se desenvolve maneira gradual, obedecendo a uma ordem sistemática de aquisição. Quatro tarefas de identificação, usando o instrumento PerceFAL, foram realizadas com 66 crianças (de ambos os gêneros), entre 4- 5 anos de idade. A tarefa consistiu na apresentação do estímulo acústico, por meio de fones de ouvido, e na escolha da gravura correspondente à palavra apresentada, dentre duas possibilidades de gravuras dispostas na tela do computador. O tempo de apresentação do estímulo e o tempo de reação das crianças foram computados automaticamente pelo software PERCEVAL. O desempenho perceptivo-auditivo de crianças ocorreu de modo gradual e é dependente da classe fônica. A maior acurácia na identificação dos contrastes fônicos parece obedecer a seguinte ordem: vogais, sonorantes, oclusivas e fricativas. O tempo de reação das respostas corretas foi menor do que das respostas incorretas (exceto para a classe das vogais). A partir dos mapas perceptuais, pode-se verificar que, no interior da classe das vogais, o parâmetro ântero-posterior parece exercer um importante papel na saliência perceptual. Para a classe das obstruintes e sonorantes (nasais e líquidas), as pistas acústicas que marcam o vozeamento (no caso das obstruentes) e o modo de articulação (no caso das sonorantes) são mais robustas perceptualmente do que as pistas de ponto de articulação. Embora a percepção da fala não deva ser reduzida a uma mera interpretação sensorial, as pistas acústicas dos segmentos da fala exercem uma importante influência para a sua categorização.*

▪ *PALAVRAS-CHAVE: Percepção de fala. Aquisição da linguagem. Fonética acústica. Português brasileiro.*

REFERENCES

ANDRÉ, C.; GHIO, A.; CAVÉ, C.; TESTON, B. **PERCEVAL**: Perception Evaluation Auditive & Visuelle (v. 5.0.30) (Programa de computador). Aix-en-Provence, 2009. Disponível em: <http://www.lpl-aix.fr/~lpldev/perceval/>. Acesso em: 23 mai. 2015.

BARTON, D. Phonemic perception in children. In: YENI-KOMSHIAN, G.; KAVANAGH, J.; FERGUSON, C. (Ed.). **Child Phonology**, v.2. New York: Academic Press, 1980. p.97-116.

BERTI, L. C. Instrumento de Avaliação da Percepção da Fala –PERCEFAL. In: **I Simpósio Internacional do Grupo de Pesquisa Avaliação da Fala e da Linguagem** – Perspectivas Interdisciplinares em Fonoaudiologia, 2011. Marília: UNESP, 2011.

BERTI, L. C.; FALAVIGNA, A. E.; SANTOS, J. B.; OLIVEIRA, R. A. Desempenho perceptivo-auditivo de crianças na identificação de contrastes fonológicos entre as oclusivas. **J. Soc. Bras. Fonoaudiol.**, São Paulo, v.24, n.4, p.348-354, 2012. Disponível em: <http://dx.doi.org/10.1590/S2179-64912012000400010>. Acesso em: 23 mai. 2015.

BERTI, L. C.; ROQUE, L. M. R. Desempenho perceptivo-auditivo de crianças na identificação de contrastes fonológicos entre as vogais tônicas. **CoDAS**, São Paulo, v.25, n.26, p.534-541, 2013.

BEST, C. Emergence of language-specific constraints in perception of non-native speech: A window on early phonological development. In: BOYSSON-BARDIES, B. et al. (Ed.). **Developmental Neurocognition**: Speech and Face Processing in the First Year of Life. Dordrecht: Kluwer Academic Publishers, 1993. p.289-304.

______. The emergence of native-language phonological influence in infants: A perceptual assimilation model. In: NUSBAUM, H.; GOODMAN, J.; HOWARD, C. (Ed.). **The Transition from Speech Sounds to Spoken Words**: The Development of Speech Perception. Cambridge, MA: MIT Press, 1994. p.167-224.

BEST, C.; McROBERTS, G. Phonological influences in infants' discrimination of two non-native speech contrasts. **Paper presented at the Society for Research in Child Development**, Kansas City, Kansas, 1989.

BORDEN, G. J.; HARRIS, K. S.; RAPHAEL, L. J. **Speech science primer**: Physiology, acoustics and perception of speech. 3. ed. Baltimore: Lippincott Williams & Wilkins, 1994.

BROWN, C.; MATTHEWS, J. The acquisition of segmental structure. In: MATTHEWS, J.; WHITE, L. (Ed.). **McGill Working Papers in Linguistics**, Special Issue on Language Acquisition, v.9. Montreal: McGill University, Department of Linguistics, 1993. p.46-76.

BROWN, C.; MATTHEWS, J. The role of feature geometry in the development of phonemic contrasts. In: HANNAHS, S. J.; YOUNG-SCHOLTEN, M. (Ed.). **Generative Studies in the Acquisition of Phonology.** Amsterdam: John Benjamins Publishing Company, 1997. p.67-112.

BROWN, C. The interrelation between speech perception and phonological acquisition from infant to adult. In: ARCHIBALD, J. (Ed.). **Second language grammars**. Oxford, England: Blackwell, 2000. p.4-63.

CASSERLY, E.; PISONI, D. B. Speech perception & production. **Wiley Interdisciplinary Reviews:** Cognitive Science, Hoboken, v.1, n.5, p.629-647, 2010.

EDWARDS, M. Perception and production in child phonology: The testing of four hypotheses. **Journal of Child Language**, Cambridge, n.2, p.205-219, 1974.

EILERS, R.; GAVIN, W.; OLLER, D. Cross-linguistic perception in infancy: early effects of linguistic experience. **Journal of Child Language**, Cambridge, n.9, p.289-302, 1982.

EIMAS, P.; SIQUELAND, E.; JUSCZYK, P.; VIGORITO, J. Speech perception in infants. **Science**, Washington, v.171, p.303-306, 1971.

FERREIRA-SILVA, A.; PACHECO, V. Evidências da relação entre duração segmental e percepção de fricativas surdas e sonoras em ataque silábico. **Confluência**, Rio de Janeiro, n.37/38, p.180-200, 2011.

GARNICA, O. The development of phonemic speech perception. In: MOORE, T. (Ed.). **Cognitive Development and the Acquisition of Language**. New York: Academic Press, 1973. p.215-222.

GATHERCOLE, S. Nonword repetition and word learning: The nature of the relationship. **Applied Psycholinguistics**, Cambridge, n.27, p.513-543, 2006.

GERRITS, E. **The categorisation of speech sounds by adults and children:** a study of the categorical perception hypothesis and the developmental weighting of acoustic speech cues. 126 f. Thesis (Doctorate degree) – Utrecht University, Utrecht, 2001.

HARDCASTLE, W. J.; LAVER, J.; GIBBON, F. E. (Ed.). **The Handbook of Phonetic Sciences**. 2. ed. Oxford: Blackwell, 2010.

HAZAN, V., BARRETT, S. The development of phoneme categorisation in children aged 6 to 12 years. **Journal of Phonetics**, London, n.28, p.377-396, 2000.

JAKOBSON, R. **Child language, aphasia and phonological universals**. The Hague: Mouton, 1968.

JOHNSON, K. **Acoustic and auditory phonetics**. London: Blackwell, 1997.

KENT, R. D.; READ, C. **The Acoustic Analysis of Speech**. San Diego, California: Singular Publishing Group, Inc., 1992.

LAMPRECHT, R. R. et al. **Aquisição fonológica do português:** perfil de desenvolvimento e subsídios para a terapia. Porto Alegre: Artmed, 2004.

LOPES, O.; CAMPOS, C. H. **Tratado de Otorrinolaringologia**. São Paulo: Editora Roca, 1994.

MACMILLAN, N. A.; CREELMAN, C. D. **Detection theory**: A user's guide. Cambridge: Cambridge University Press, 1991.

MEHLER, J. Language related dispositions in early infancy. In: MEHLER, J.; FOX, R. (Ed.). **Neonate Cognition**: Beyond the Blooming, Buzzing Confusion. Hillside, NJ: Erlbaum, 1985.

MILLER, G. A.; NICELY, P. E. An analysis of perceptual confusions among some English consonants. **Journal of the Acoustical Society of America**, Melville, n.27, p.338-352, 1955.

MOTA, H. B. Pares mínimos: os contrastes do português brasileiro. **Pró-Fono**, Barueri, n.13, p.98-106, 2001.

MUNSON, B.; EDWARDS, J.; BECKMAN, M. E. Relationships between nonword repetition accuracy and other measures of linguistic development in children with phonological disorders. **Journal of Speech, Language, and Hearing Research**, Rockville, n.48, p.61-78, 2005.

PANNETON, R.; NEWMAN, R. Development of Speech Perception. In: WENER, L.; FAY, R.; POPPER, A. N. (Ed.). **Human Auditory Development**. New York, USA: Springer New York Dordrecht Heidelberg London, 2011. p.197-222.

PATER, J.; STAGER, C.; WERKER, J. The Perceptual Acquisition of Phonological Contrasts. **Language**, Washington, n.80, p.384-402, 2004.

PISONI, D. B.; TASH, J. Reaction times to comparisons within and across phonetic categories. **Perception & Psychophysics**, Austin, n.15, v.2, p.285-290, 1974.

POLKA, L.; WERKER, J. F. Developmental changes in perception of nonnative vowel contrasts. **Journal of Experimental Psychology**: Human Perception and Performance, Washington, n.20, v.2, p.421-435, 1994.

POLKA, L.; BOHN, OS. A cross-language comparison of vowel perception in English-learning and German-learning infants. **The Journal of the Acoustical Society of America (JASA)**, Melville, n.100, p.577-592, 1996.

SHVACHKIN, N. Kh. The development of phonemic speech perception in early childhood. Traduzido por E. Dernbach e republicado em 1973. In: FERGUSON, C.; SLOBIN, D. (Ed.). **Studies of Child Language Development**. New York: Holt, Rinehart and Winston, 1948. p.91-127.

STREETER, L. Language perception of 2-month-old infants shows effects of both innate mechanisms and experience. **Nature**, London, v.259, p.39-41, 1976.

TREHUB, S. The discrimination of foreign speech contrasts by infants and adults. **Child Development**, Ann Arbor, v.47, p.466-472, 1976.

TENANI, L.; SILVEIRA, A. A. M. D. O alçamento das vogais médias na variedade culta do noroeste paulista. **ALFA:** Revista de Linguística, v.52(2), p.467-464, 2008.

WANG, M. D.; BILGER, R. C. Consonant confusion in noise: a study of perceptual features. **The Journal of the Acoustical Society of America (JASA)**, Melville, v.54, p.1248-1266, 1973.

WERKER, J.; GILBERT, J.; HUMPHREY, K.; TEES, R. Developmental aspects of crosslanguage speech perception. **Child Development**, Ann Arbor, v.52, p.349-353, 1981.

WERKER, J.; TEES, R. Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. **Infant Behavior and Development**, Amsterdam, v.7, p.49-63, 1984a.

_______. Phonemic and phonetic factors in adult cross-language speech perception. **The Journal of the Acoustical Society of America (JASA)**, Melville, v.75, p.1866-1878, 1984b.

_______. Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. **Infant Behavior & Development**, Amsterdam, v.25, p.121-133, 2002.

WERKER, J.; LALONDE, C. Cross-language speech perception: Initial capabilities and developmental change. **Developmental Psychology**, Washington, v.24, p.672-683, 1988.

WERKER, J.; POLKA, L. Developmental change in speech perception: New challenges and new directions. **Journal of Phonetics**, London, v.21, p.83-101, 1993.

WINTERS, S. Turning phonology inside out: testing the relative salience of audio and visual cues for place of articulation. In: LEVINE, R.; MILLER-OCKHUIZEN, A.; GONSALVEZ, A. J. (Ed.). **Ohio State Working Papers in Linguistics**, n.53, p.168-199, 2000.

Received in june 2015

Approved in august 2016