# Meaning Does Not Affect Consonant Discrimination Accuracy or Response Time in A Same-Different Segment Comparison Task

## *O significado não afeta a precisão da discriminação de consoantes ou o tempo de resposta em uma tarefa igual-diferente de comparação de segmentos*

Rui Rothe-Neves

Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Minas Gerais / Brasil
rothe-neves@ufmg.br
http://orcid.org/0000-0002-8896-8862

Cibelle de Mesquita Duarte

Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Minas Gerais / Brasil
cibellemd@yahoo.com.br
https://orcid.org/ /0000-0002-0988-926X

**Abstract:** This article reports on three studies designed to test whether knowing the meaning of a word can influence the ability to discriminate sounds in it. In a same-different paradigm that required overt segmentation, we investigate the ability to compare consonants in the onset position of a pair of one-syllable pseudowords before (pre-test) and after (post-test) a training phase in which we attributed meanings to half of the pseudowords used. Reduced response time and increased accuracy (percentage of correct answers) in the post-tests revealed a training effect in two experiments. Still, there was no difference between pseudowords to which meanings were attributed or not. Conclusion: Knowing the meaning of a word does not influence the ability to discriminate sounds in it.

**Keywords:** speech perception; phoneme discrimination; segmentation; psycholinguistics.

**Resumo:** Este artigo relata dois experimentos planejados para testar se o conhecimento do significado de uma palavra pode influenciar a capacidade de discriminar sons nela presentes. Em um paradigma igual-diferente que requer segmentação explícita, investigou-se a capacidade de comparar consoantes na posição de início em uma pseudopalavra monossilábica antes (pré-teste) e depois (pós-teste) de uma fase de treinamento em que se atribuíram significados a metade das pseudopalavras utilizadas. Em ambos os experimentos, verificou-se efeito de treino na comparação de segmentos com tempo de resposta reduzido e maior precisão (porcentagem de respostas corretas). Ainda assim, não houve diferença entre as pseudopalavras às quais foram atribuídos significados ou não. Conclusão: Saber o significado de uma palavra não influencia na habilidade de discriminar os sons nela contidos.

**Palavras-chave:** percepção de fala; discriminação de fonemas; segmentação; psicolinguística.

# 1 Theoretical Motivation

Lexical access occurs when the meanings of words are contacted in long-term semantic memory. Speech perception precedes lexical access and comprises the "mapping of the highly variable acoustic signal to a linguistic representation" (HOLT; LOTTO, 2010). There are two types of theoretical formulations about what goes on in the speaker's mind when s/he hears a word and before s/he knows if it belongs to his language. Interactive models (for example, McCLELLAND; ELMAN, 1986) assume the interaction between linguistic knowledge and the early stages of auditory evaluation. On the other hand, non-interactive models (for example, NORRIS; MCQUEEN, CUTLER, 2000; see also PISONI; LUCE, 1987) assume that linguistic knowledge does not influence the first stages of acoustic-phonetic analysis after auditory transduction. This would be a "linguistically naïve" stage of auditory evaluation, and only afterward does the influence of linguistic knowledge occur. According to non-interactive models, the speaker's knowledge of his language would be necessary not to extract those linguistically essential components from the acoustic signal but only for "category labeling", that is, to know whether such a sound belongs to the category [s] or [z] (for a review, see KINGSTON 2009).

The state-of-affairs just described is what Eysenck and Keane (2005) presented as the two approaches to auditory stimulus processing. The first is serial processing, in which only one process occurs at a time, and processing is complete before another process starts. Some consider such as a simplified approach since it only considers bottom-up processing, ignoring the influence of the individual's knowledge and expectations, known as top-down processing. On the other hand, parallel processing argues that two or more processes can run simultaneously. A common form of processing is known as cascade processing, in which later processes start before some of the previous processes finish.

One source of evidence in favor of interactive models is the so-called lexical identification shift (GANONG, 1980). This effect consists of shifting the phoneme boundary in forced-choice identification tasks towards the one end of the continuum where the stimuli form a word instead of a nonword. However, we already know that in the identification tasks, the participant is biased towards stored knowledge because he hears a stimulus (sound, syllable, or word) and must identify it as a sound from his language stored in memory (SCHOUTEN; GERRITS; VAN HESSEN, 2003). It is, therefore, a task in which category labeling and linguistic knowledge have a predominant role. Kingston (2009) used a same-different discrimination paradigm, in which acoustic information is predominant. A "same-different" task requires low levels of auditory memory load, but it cannot alone absolutely prevent participants from showing labeling behavior (GERRITS; SCHOUTEN, 2004). Burton, Small and Blumstein (2000) used a discrimination task in the same-different paradigm in which participants were required to pursue an overt segmentation, an idea we develop further here. The authors concluded that lexical effects are postperceptual.

The present research aimed to investigate whether knowledge about the meaning of a word can influence the ability to discriminate sounds in it. We used a pre-and post-test design, with an in-between phase when participants learned the meanings of half the pseudowords used in a discrimination task. Although there is a long tradition of research on the effect of meaning on auditory word recognition or speech perception research (as far as we can see, inaugurated by SPREEN; BORKOWSKI; BENTON, 1967), to the best of our knowledge, this research strategy has only been used on reading (WHITTLESEA; CANTWELL, 1987).

## 2 Selection of test items

In this section, we report how we selected the best set of items for the task at hand, as described below.

### 2.1 Methods

Forty-seven male and female adults (18-62 years, mean: 32 years) without self-reported hearing problems participated in the study after informed consent. Seven participants were excluded due to hearing loss or inadequate task performance.

The segment comparison task is a discrimination task in a same-different paradigm with overt segmentation. The task was based on the study by Silva (2007), which, in turn, used the general idea of a similar task (BURTON; SMALL; BLUMSTEIN, 2000). The stimuli were the same as for the discrimination task described in Rothe-Neves, Lapate e Pinto (2004). In each trial, the participants heard a pair of CVC-syllables and were required to make the same/different judgment about phonetic segments. The consonants at onset were the target segments and differed by manner, place, or voice. The remaining VC segments were always different.

Consequently, participants had to segment out the pair's initial consonants and compare them. Only then could they make the discrimination judgment. By segmentation, we refer to the process whereby a participant separates the individual segments from the word stimulus to complete the task. (BURTON; SMALL; BLUMSTEIN, 2000, p. 680).

The task was composed of pairs of heavy syllables that bear no meaning in Portuguese. Heavy syllables have a coda, and, in Portuguese, the coda can be one of the variants of /l/ (velarized or vocalized), of /s/ (alveolar or alveopalatal), of /r/ (with several possible phonetic manifestations, including none), a glide, or the nasal /N/, which typically does not show up phonetically besides nasalizing the preceding vowel. All Portuguese consonant and vowel phonemes were combined according to the phonotactic restrictions, and excluding actual words (e.g., *mar* "sea"). In some pairs, the onset of the syllable (the initial consonant) was the same in both syllables ("jon" [ʒõ], "jar" [ʒax]). In the other pairs, the consonants differed in terms of place of articulation, manner of articulation, or voicing ("xon" [ʃõ], "jar" [ʒax]). The rhyme was never the same between syllables of the pair. Thus, this manipulation resulted in 196 pairs of syllables.

Data collection consisted of a hearing screening and the segment comparison task. The evaluations were carried out in a quiet room of

a private clinic. Possible hearing loss was excluded through a hearing screening consisting of pure tone audiometry at frequencies 500Hz, 1000Hz, 2000Hz, and 4000Hz in the right and left ears separately with an adequately calibrated Amplaid A177 audiometer. The normality reference value for tonal hearing threshold was up to 25 dB hearing level (LLOYD; KAPLAN, 1978). We used the software PercEval (Université de Provence / CNRS, Brazilian version: UFMG Phonetics Laboratory) installed on an HP G42-240BR notebook and a Leadership headset. The participants heard the stimuli binaurally and used the keyboard to respond whether the first consonant in each syllable of the pair was the same or different. The following instructions were presented: "You will hear two syllables each time and must decide whether the syllables begin with the same sound. If you think so, press the [S] key. If you do not think so, press the [N] key. If you do not know, choose any alternative and respond as quickly as you can".

Each trial consisted of a pair of syllables, as described, separated by 300 ms interval and followed by up to 3 seconds when the participants should respond. After a 0.5 s pause following the response, the subsequent trial was presented with a randomly selected syllable pair. The response variable is it the percent correct score.

## 2.2 Results

The analysis's first step was to exclude items with percent correct answers below 20% and above 80% correct. Then, the reliability coefficient was estimated. It is a measure of the consistency with which a set of items evaluates a characteristic (here the ability to compare mental sound units abstracted from the context in which they appeared). It is expressed by Cronbach's alpha coefficient – in this case, estimated by the Kuder-Richardson method for dichotomous responses (yes/no). For statistical analyzes, we used SPSS, version 12.

Table 1 – Items selected for the test studies and corresponding statistics

| Items | Scale Mean if item excluded | Scale Variance if item excluded | Corrected item-total correlation | Cronbach Alpha if item excluded |
|---|---|---|---|---|
| SAR-ZIR | 17,895 | 20,745 | ,484 | ,849 |
| FER-VES | 17,737 | 21,280 | ,430 | ,851 |
| JÃO-XUN | 17,711 | 21,346 | ,434 | ,850 |
| FIR-VES | 17,737 | 20,848 | ,544 | ,847 |
| JÃO-ZOR | 17,711 | 20,806 | ,583 | ,845 |
| SAR-ZEI | 17,921 | 20,561 | ,521 | ,847 |
| SIR-ZAI | 17,789 | 20,711 | ,539 | ,847 |
| SIR-ZEI | 17,816 | 19,722 | ,772 | ,837 |
| SOR-ZAI | 17,763 | 20,942 | ,498 | ,848 |
| SOR-ZEI | 17,737 | 21,064 | ,487 | ,849 |
| JOS-XUN | 17,763 | 20,672 | ,568 | ,846 |
| TAN-TEI | 17,711 | 21,130 | ,494 | ,848 |
| TAN-DEM | 17,763 | 20,726 | ,554 | ,846 |
| MUS-MON | 17,605 | 22,570 | ,172 | ,858 |
| BES-BUR | 17,684 | 21,898 | ,306 | ,855 |
| XUN-XIR | 17,684 | 21,681 | 21,681 | ,853 |
| TOU-TUR | 17,658 | 22,934 | ,030 | ,862 |
| KEU-KAR | 17,947 | 22,321 | ,126 | ,863 |
| JÃO-JER | 17,684 | 21,465 | ,427 | ,851 |
| BIR-BUS | 17,632 | 23,374 | -,095 | ,865 |
| GUS-GOR | 17,632 | 22,293 | ,237 | ,856 |
| NUI-NEU | 17,658 | 21,637 | ,408 | ,851 |
| PUR-PEI | 17,632 | 21,212 | ,585 | ,847 |
| ZAI-ZIR | 17,632 | 22,077 | ,305 | ,854 |

Source: Created by the authors.

Out of a possible total of 7840 data points (196 pairs rated by 40 participants), ten responses were not recorded either due to exceeding the 3s-time for registration or pressing a wrong key. Twelve syllable pairs with correct answers between 20-80% (average accuracy of 73%) form the set based on the reliability index that best discriminates the research participants' ability to perform the task. We then included 12 items next in the difficulty scale: the 12 most accessible items after the ones already selected. Table 1 shows the 24 items (average accuracy = 76,7%; α = 0.86) selected for the studies to follow.

## 3 Segment comparison – Experiment 1

### 3.1 Methods

Twenty-one adult participants (18-27 years, average: 20.5 years) without a history of hearing problems performed the segment comparison task on three occasions (pre-test and two post-tests). The task was the same as in the first study, except that the participants heard the stimuli with white noise added in a 0dB signal-to-noise ratio (SNR) to make it more difficult to perceive the contrast under test and increase response variability (SANTOS; LEMOS; ROTHE-NEVES, 2014).

After the pre-test, participants took home a list of random meanings for those 12 items that formed the previous study's best set. In each pair of pseudowords, only one received a meaning. Participants had to learn or memorize the meanings in the list. After five days, participants performed a learning verification task, which consisted of listening to one of the pseudowords in the headset and simultaneously viewing one of the meanings presented on a computer screen. Participants should answer whether the meaning corresponded to the word they heard or not using the keyboard. Half the pseudowords in the task had a corresponding meaning, with an expected "yes" answer, and for the other half, the answer should be "no." Only eight participants answered correctly at least 80% of the time and had their results further included in this study. Participants performed the first post-test after the verification task and a second post-test after two days to investigate a persistent effect. The collected measures were the proportion of correct answers (accuracy) and each participant's response time. A mixed-effects design with Session (Pre-test; Post-test 1; Post-test 2) and Condition (meaning; no meaning) as repeated-measures factors, and Subject as a random effect compared the results. If bearing a meaning facilitates comparing segments, the accuracy would be higher and the reaction time lower, but we would expect no such improvement for the no-meaning condition. The analyses were generalized linear models of the binomial type with a logit link function for the accuracy results and variance analysis for the response time; both run in *lme4* package for R (BATES et al., 2015; R CORE TEAM, 2021). A maximum likelihood test comparing the full model to a model with that effect omitted determined each effect's statistical significance.

## 3.2 Results

In all, data from eight participants were analyzed, with 569 valid responses (eight subjects X three sessions X 24 items, minus seven non-computed responses), in the pre-test and two post-tests. Accuracy results (Figure 1) improved in the meaning condition and the no-meaning condition in both post-tests. In the no-meaning condition, we observe a more notable variation in the results. Participants obtained 58.51% of correct answers in the pre-test, which increased to 75% in the first post-test and 88.42% in the second post-test. In the meaning condition, the percentage of correct answers also improved from the pre-test (67.7%) to the first post-test (75.3%) and persisted to the second post-test (86.3%).

A learning effect was observed from one session to another, as shown by the higher accuracy rate. There was a significant main effect of Session ($\chi^2$ (2) = 30.9, p < .0001) and no effect of Condition ($\chi^2$ (3) = 1.9, p = .58). However, a post hoc comparison showed that the only significant difference was between the pre- and the second post-test in the no-meaning condition (p = 0.003).

Figure 1 – Accuracy in the segment comparison task as a function of Condition and Session (Study 2)
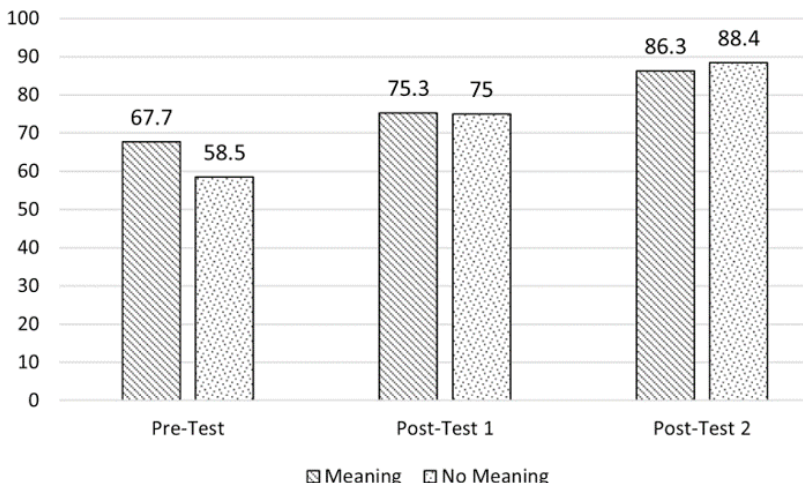


Source: Created by the authors.

The likelihood ratio (LR) test assesses the fit of a generalized linear model of the binomial type by comparing the residuals' deviation from the model to the deviation of a so-called "null model". The model is neutral or null, as it only includes the intercept and no linguistic variables as explanatory. Compared to a null model, the model presented here is significant (LR = 32.9, p < .0001).

Figure 2 – Response time (in miliseconds) in the segment comparison task as a function of Condition and Session (Study 2)
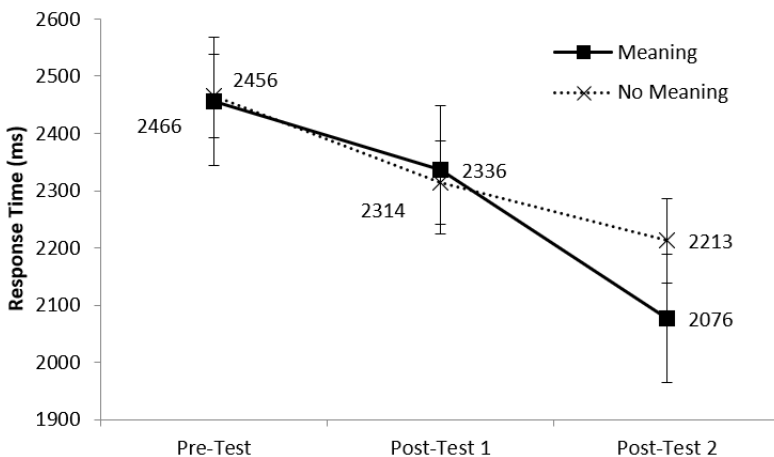


Source: Created by the authors.

The response time was reduced from the pre-test to the post-tests both in the meaning and no-meaning conditions. The average time was 2456.5 ms in the pre-test, 2336.4 ms in the first post-test, and 2076.4 ms in the second post-test in the meaning condition. In the no-meaning condition, the reaction time was 2465.51 ms in the pre-test, 2314.07 ms in the first post-test, and 2212.69 ms in the second post-test. For the ANOVA, the response time was log-transformed. Again, there was a significant main effect of Session ($F[2,563] = 10.4$, $p < .0001$), but no effect of Condition ($F[3,563] = 0.54$, $p = 0.65$). A likelihood-ratio test confirmed that the model is significant as compared to a null model (LR = 18.6, $p < .002$). Post hoc tests confirmed that the only significant difference was between the pre- and the second post-test in the meaning condition ($p = 0.0005$).

In sum, participants' accuracy showed an effect in the no-meaning condition between the first and the last session. On the other hand, reaction times showed an effect in the meaning condition. As it is impossible to obtain a result favourable to the hypothesis of the influence of meaning in one measure and an opposite result in the other, we interpreted the results of the second post-test as an apparent learning effect due to the participants' experience with the task itself. However, there may have been a possible alternative hypothesis that we consider next.

The best stimulus set identified in the previous section and used in the meaning condition ended up almost all with different-consonant pairs, and the same-consonant pairs were in the no-meaning condition. Except one, all pairs in the meaning condition begin with consonants that differ in voice. On the other hand, in the 12 pairs included in the "no meaning" condition, all initial consonants are the same, and again we have a single exception. Remember that the task asks if the initial consonants in each syllable in the pair are the same. As these are natural stimuli registered by a speaker for the task, each acoustic wave is slightly different, irrespective of whether these different sounds belong to the same linguistic categories. Therefore, when comparing the initial consonant of the first pair with that of the second pair, one would expect the listener to have more difficulty if the sounds are linguistically the same (but acoustically different) than if they are linguistically different. Then, it seems convenient to replicate the study with a more balanced set of pseudowords.

## 4 Segment comparison – Study 3

### 4.1 Methods

Twenty-eight adult participants (17-57 years, average: 28 years) without a history of hearing problems performed the same segment comparison task before. Fourteen participants did not continue in the study due to inadequate performance. Except for a different choice of which pseudowords would "acquire" meaning, the segment comparison and the learning tasks, materials, and procedure were the same as in Study 2. After the pre-test, the same random meanings were assigned to 12 randomly selected syllables, six from each set. In this way, the syllables

were counterbalanced concerning their initial status. As two post-tests seemed not to contribute to the question at hand, in this experiment, we compared the results in the pre-test with only one post-test.

## 4.2 Results

In all, fourteen participants provided 671 data points for analysis in the pre-test and post-test. Figure 3 shows a slight improvement in the percentage of correct pseudowords, which acquired meaning, from 82.7% in the pre-test to 83.3%. On the other hand, in the no-meaning condition, the correct percentage increased from 73.8% in the pre-test to 84.4%. It is, thus, not possible to point out a possible effect of initial difficulty caused by the phonological structure of the syllables, contrary to what was the case in the previous study.

Figure 3 – Accuracy in the segment comparison task as a function of Condition and Session (Study 3)
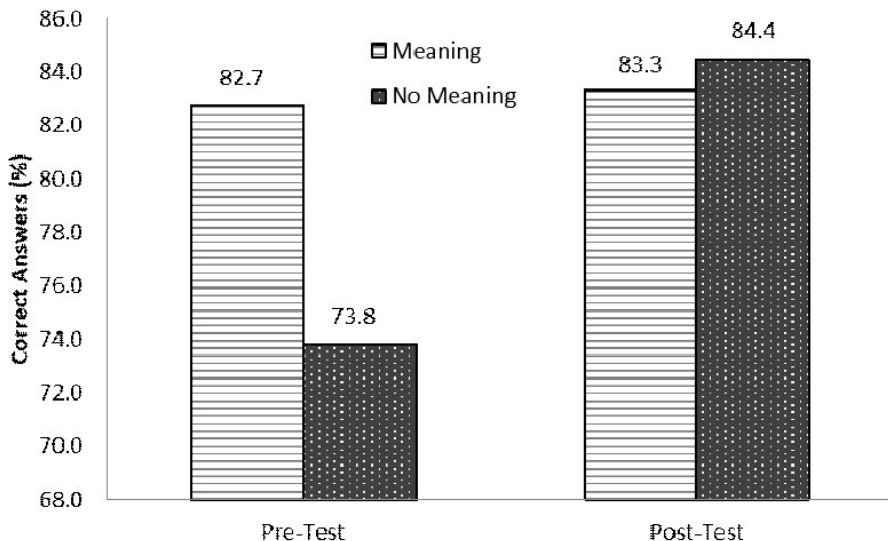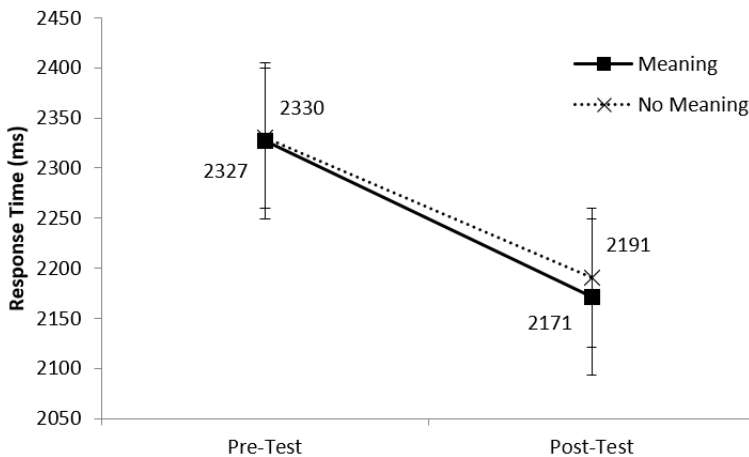


Source: Created by the authors.

There was no effect of Session ($\chi^2$ (1) = 3,44; p = 0.06) or Condition (meaning versus no meaning) remains nonsignificant ($\chi^2$ (2) = 4,03; p = 0.13). As a result, the null model without linguistic information is closer to the whole model, as the "linguistic information" was not significant; the model fit is therefore very modest (LR = 7.48; p = 0.058). Even so, the difference between the two is close to statistical significance. A post hoc comparison revealed that a (barely) significant difference was between the two conditions in the pre-test (p = 0.0485).

Here we report response time results slightly differently than in the previous section. The data resulted in an average response time of 2282 ms (sd = 699.75 ms). In a linear mixed-effects model, as in the previous study, the independent variables were Condition (meaning; no meaning) crossed with Session (Pre-test; Post-test). However, the residue deviations' analysis revealed 12 observations that resulted below or above 2.5 standard deviations. Such observations were not due to session, participant, condition, or item. Thus, we report the response time results without these possibly marginal observations, which represented only 1.79% of the data.

Figure 4 – Response time (in miliseconds) in the segment comparison task as a function of Condition and Session (Study 3)



Source: Created by the authors

Figure 4 shows the effect of training or learning on the task. For those pseudowords with meaning, response time decreased from 2377 ms

in the pre-test to 2191 ms, and in the no-meaning condition, from 2354 ms to 2207 ms. The main effect of Session was significant ($F[1,655] = 18.54$; $p = 0.0019$), but again, there was no effect of Condition ($F[2,655] = 0.02$; $p = 0.98$).

As in the previous study, we again observed a progressive improvement of correct answers between sessions in both the meaning and the no-meaning conditions, a learning effect between sessions, and a reduction of reaction time.

## 5 Discussion

Both segment comparison studies reported here showed similar results despite the difference in participants and pseudowords. There was also a learning effect in the second post-test, even if it happened five days after the first post-test.

We found an effect from the pre-test to the second post-test on accuracy results in the no-meaning condition and response times in the meaning condition. Contrary to Whittlesea and Cantwell (1987), we found no evidence that knowledge about the meaning of words influence the ability to discriminate consonants at syllable onset from the pre-test to the post-test in both experiments. We found more variation in the results in the no-meaning condition, reinforcing the conviction that meaning does not influence the learning effect we detected in comparing segments.

The task we used is known to reduce response bias towards the stored representations in long-term memory. Nonetheless, the overt segmentation required participants to abstract away from the basic acoustic features. Because the phonetic context was always different, coarticulation causes a segment's production with different phonetic details. The participants compared sounds that were not the same, although their phonetic category (e.g., [ʒ]) is. Thus, we might safely say that the task taps onto the result of acoustic-phonetic processing, not at the bottom-level acoustic sound.

As for the interpretation of the results in terms of a cognitive architecture, the research presented here is compatible with the view that an autonomous level accomplishes speech perception, a level lexical knowledge does not affect (BURTON, SMALL; BLUMSTEIN, 2000; KINGSTON, 2009).

## References

BATES, D.; MÄCHLER, M.; BOLKER, B.; WALKER, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48. doi:10.18637/jss.v067.i01.

BURTON, Martha W.; SMALL, Steven L.; BLUMSTEIN, Sheila E. The role of segmentation in phonological processing: an fMRI investigation. *Journal of cognitive neuroscience*, v. 12, n. 4, p. 679-690, 2000.

EYSENCK, Michael W.; KEANE, Mark T. *Cognitive psychology*: A student's handbook. Taylor & Francis, 2005.

GANONG, William F. Phonetic categorization in auditory word perception. *Journal of experimental psychology: Human perception and performance*, v. 6, n. 1, p. 110, 1980.

GERRITS, Ellen; SCHOUTEN, M. E. H. Categorical perception depends on the discrimination task. *Perception & psychophysics*, v. 66, n. 3, p. 363-376, 2004.

HOLT, Lori L.; LOTTO, Andrew J. Speech perception as categorization. *Attention, Perception, & Psychophysics*, v. 72, n. 5, p. 1218-1227, 2010.

KINGSTON, John. Ears to categories: New arguments for autonomy. In: FROTA, Sonia; VIGÁRIO, Marina; FREITAS, Maria João (Ed.). *Prosodies*: With special reference to Iberian languages. Walter de Gruyter, 2009. p. 177-222.

MCCLELLAND, James L.; ELMAN, Jeffrey L. The TRACE model of speech perception. *Cognitive psychology*, v. 18, n. 1, p. 1-86, 1986.

NORRIS, Dennis; MCQUEEN, James M.; CUTLER, Anne. Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences*, v. 23, p. 299-325, 2000.

PISONI, David B.; LUCE, Paul A. Acoustic-phonetic representations in word recognition. *Cognition*, v. 25, n. 1-2, p. 21-52, 1987.

R CORE TEAM (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Available at https://www.R-project.org/.

ROTHE-NEVES R., LAPATE R..C.; PINTO J. S. Tarefa de discriminação de fonemas com pseudopalavras. Revista de Estudos da Linguagem. Faculdade de Letras da UFMG, v. 12, n. 2, 159-179. jul.-dez. 2004.

SANTOS, Lílian Marinho; LEMOS, Stela Maris Aguiar; ROTHE-NEVES, Rui. Perceptual confusions among consonants in Brazilian Portuguese as a function of noise. *Audiology-Communication Research*, v. 19, n. 2, p. 145-152, 2014.

SCHOUTEN, Bert; GERRITS, Ellen; VAN HESSEN, Arjan. The end of categorical perception as we know it. *Speech communication*, v. 41, n. 1, p. 71-80, 2003.

SPREEN, Otfried; BORKOWSKI, John G.; BENTON, Arthur L. Auditory word recognition as a function of meaningfulness, abstractness and phonetic structure. *Journal of Verbal Learning and Verbal Behavior*, v. 6, n. 1, p. 101-104, 1967.

SILVA, F. G. *Estudo correlacional entre o desempenho em tarefas lingüísticas e audiológicas de indivíduos afásicos*. 2007. 92 f. Dissertação (Mestrado em Estudos Lingüísticos) – Faculdade de Letras, Universidade federal de Minas Gerais, Belo Horizonte, 2007.

WHITTLESEA, Bruce WA; CANTWELL, Alisa L. Enduring influence of the purpose of experiences: Encoding-retrieval interactions in word and pseudoword perception. *Memory & Cognition*, v. 15, n. 6, p. 465-472, 1987.