



## IMMEDIATE AND LONG-TERM EFFECTS OF A PERCEPTUAL TRAINING IN THE PRODUCTION OF WORD-INITIAL /S/-CLUSTERS IN BRAZILIAN PORTUGUESE/ENGLISH INTERPHONOLOGY

### EFEITOS IMEDIATOS E DE LONGO PRAZO DE UM TREINAMENTO PERCEPTUAL DE ENCONTROS CONSONANTAIS INICIADOS POR /S/ NA INTERFONOLOGIA DO PORTUGUES DO BRASIL/INGLES

Melissa Bettoni<sup>1</sup>

**Resumo:** O presente estudo teve como objetivo investigar os efeitos do treinamento perceptual na produção de encontros consonantais iniciados por /s/ na interfonologia do português brasileiro/inglês. As hipóteses propunham que haveria melhoria na produção após o treinamento e que tal melhoria também seria encontrada em um teste de retenção após oito meses e em um teste de retenção após doze anos. Além disso, foi levantada a hipótese de que a produção de /s/+soantes permaneceria mais problemática do que a produção de /s/+plosivas após o treinamento. Vinte e três brasileiros que falavam inglês como língua estrangeira com um nível mínimo de proficiência B1 participaram do estudo. Oito deles foram designados a um grupo de controle. Os quinze participantes restantes fizeram um teste que consistiu em tarefas de leitura e uma entrevista antes e depois de um treinamento perceptivo de identificação seguindo uma abordagem de alta variabilidade. Oito meses depois, oito participantes do grupo experimental fizeram um teste de retenção com as mesmas tarefas dos testes anterior e posterior ao treinamento perceptual. Doze anos depois, uma participante fez o teste de retenção novamente. Os resultados indicaram que a produção de encontros consonantais iniciados por /s/ no começo de palavras melhorou significativamente logo após o treinamento e que a melhora também foi encontrada nos testes de retenção de oito meses e de doze anos, mesmo para encontros que não haviam sido treinados. Os resultados indicaram que a produção de /s/+soantes continuou a ser mais problemática do que a de /s/+plosivas e que vozeamento foi o desvio de pronúncia mais persistente na produção dos encontros iniciados por /s/. A transferência da melhora da percepção para a produção pode indicar que existe uma representação mental comum subjacente a ambos os domínios (perceptual e de produção) e que as mudanças no desempenho na produção podem refletir mudanças na representação mental de sons não-nativos.

**Palavras-chave:** Treinamento Perceptual. Interfonologia. Ensino e Aprendizagem de Inglês.

**Abstract:** The present study aimed at investigating effects of perceptual training in the production of word-initial /s/-clusters in Brazilian Portuguese/English interphonology. The hypotheses proposed that there would be improvement in production after training and that such improvement would also be found in an eight-month follow-up test and in a twelve-year follow-up test. Also, it was hypothesized that the production of /s/+sonorants would remain more problematic than the production of /s/+stops after training. Twenty-three Brazilians who spoke English as a foreign language at a minimum B1 level of proficiency participated in the study. Eight of them were assigned to a control group. The fifteen remaining participants took a test consisting of reading tasks and an interview before and after an identification perceptual training following a high variability approach. Eight months later, eight participants from the experimental group took a follow-up test with the same tasks from pre- and posttests. Twelve years later one participant took another follow-up test. Results indicated that production of word-initial /s/-clusters improved significantly right after training and that the improvement was also found in both eight-month and twelve-year follow-

---

<sup>1</sup> Instituto Federal de Santa Catarina (IFSC), Chapecó, SC, Brasil e Universidade Federal da Fronteira Sul (UFFS), Chapecó, SC, Brasil.. [mebettoni@gmail.com](mailto:mebettoni@gmail.com)  
Orcid: <https://orcid.org/0000-0002-2522-1083>

up tests even for clusters which had not been trained. /s/+sonorants remained more problematic than /s/+stops and voicing was found to be more persistent than other mispronunciations of word-initial /s/-clusters. The transfer of improvement from perception to production may indicate that there is a common mental representation underlying both domains and that changes in production performance may reflect changes in mental representation of non-native targets.

Keywords: Perceptual training. Interphonology. English teaching and learning.

## 1. INTRODUCTION

For over thirty years, research has been carried out regarding perceptual training of non-native contrasts (e.g., STRANGE; DITTMAN, 1984; JAMIESON; MOROSAN, 1986; FLEGE, 1989; LIVELY; LOGAN; PISONI, 1993; BRADLOW; PISONI; AKAHANE-YAMADA; TOKURA, 1997; HARDISON, 2000, 2003; TRAPP; BOHN, 2000; YEON, 2004; BETTONI; KOERICH, 2009; COOKE; GARCIA LECUMBERRI, 2018; ALVES; KAMPFF, 2019; KABAKOFF; GO; LEVI, 2020). Perceptual training studies usually consist of exposing participants to auditory stimuli and requiring them to identify the stimulus heard as one out of two stimuli represented visually (e.g., letters, sounds, words, symbols, pictures) on a screen. Immediate feedback is usually provided with repetition of the stimuli when the incorrect answer is given. The perceptual training is often comprised of several sessions of variable duration lasting from days to weeks with a pretest being answered before training and a posttest right after the last session is carried out. Few studies have reported lasting results after months or years with a retention test (e.g., BETTONI; KOERICH, 2009; ALVES; KAMPFF, 2019).

One of the main goals of perceptual training studies is generalization of improvement to other perceptual tasks, voices, sounds and to production as well. Rochet (1995) tested transfer of perceptual training of twelve Mandarin speakers on the contrast between French /pu/ vs. /bu/ to production with synthesized stimuli manipulating the voice onset time (VOT) of the consonants. Besides the perceptual training, participants had an imitation component differentiating it from the usual perceptual training only procedure. However, no instruction on production was provided.

Several other studies have also found evidence of improvement in production with synthesized and natural stimuli (e.g.; TRAPP; BOHN, 2000; LAMBACHER et al., 2005). Since transfer to production has occurred without explicit instruction on articulation of sounds and production in general, Bradlow and colleagues (1997, p. 308) concluded that “there is a unified, common mental representation that underlies both speech perception and production”. Production performance can, thus, reflect changes in the phonological mental representations of non-native categories.

The present study investigated effects of perceptual training in the production of word-initial /s/-clusters in Brazilian Portuguese/English interphonology. Word-initial /s/-clusters were chosen as the target of the study because research has shown that mere exposure to native productions of /s/-clusters as well as one-time explicit instruction have not triggered target-like productions. Mispronunciation of word-initial /s/-clusters in English may be produced by Brazilian Portuguese speakers of English even when they are proficient in English and have lived and worked for years in a country where English is a native language and have used English daily as their main language (BETTONI-TECHIO, 2008). One of the reasons is transfer from Brazilian Portuguese phonotactics which trigger prothesis, palatalization, and voicing of the word-initial /s/ in /s/-clusters. According to Yildiz (2005), word-initial /s/-clusters are stored as a complex segment and thus they are more difficult to be acquired by both L1 and L2 speakers than non-/s/-clusters.

The present study was carried out considering the persistent failure in producing target-like word-initial /s/-clusters in Brazilian Portuguese/English interphonology and the successful studies in which perceptual training has been administered to promote target-like production of non-native contrasts (e.g., YEON, 2004; ALVES; KAMPFF, 2019; among many others). It aimed at investigating effects of a perceptual training in the production of word-initial /s/-clusters in Brazilian Portuguese/English interphonology. Based on the studies which have been carried out on perceptual training and word-initial /s/-clusters, the hypotheses set up were the following.

Hypothesis 1: There will be improvement in production from the pretest to the posttest.

Hypothesis 2: /s/+sonorant clusters will remain more problematic than /s/+stop clusters after training.

Hypothesis 3: There will be retention of improvement in production after eight months.

Hypothesis 4: There will be retention of improvement in production after twelve years.

## **2. METHOD**

### **2.1 Participants**

The participants of the study were 23 Brazilian speakers of English. Their experience with English varied from level B1 to C2 in the Common European Framework Reference for Language. Fifteen participants were assigned to the experimental group participating in the perceptual training and answering pre and posttests. Eight participants were assigned to the control group and answered both pre and posttest without participating in the training program. At the time of the training program, in the experimental group, six participants were adult males with ages ranging from 16 to 31 and seven were adult females with ages ranging from 20 to 55. The two remaining participants were pre-adolescent girls with ages 9 and 11 – P1 and P10 respectively. The pre-adolescents were grouped with the adults because it was observed that they differed between them in terms of performance in a way similar to the adult participants, and they improved as much as the adults as well. P1 had six years and P10 had seven years of previous formal instruction of English in a language school. Nine out of fifteen participants in the experimental group were either undergraduate Letras students or had graduated in Letras and already worked as English teachers. Considering the other six participants one had studied English previously for two years and the other five from six to ten years. Out of the fifteen participants in the experimental group, eight took the eight-month retention test and one took the twelve-year retention test. The participant who took the twelve-year retention test was P1, the youngest participant of the study. She was nine when the first data gathering was carried out and twenty-one years old when the final retention test was taken.

### **2.2 Data gathering instruments**

The first instrument was a background questionnaire aimed at assessing biographical information which could be used to control the variables in the study such as age, sex, amount of exposure to English, and knowledge of other languages. In order to assess production before and after training, five other instruments were designed: a reading task in Brazilian Portuguese, a short semi-guided interview, a paragraph-reading task, a story-reading task, and a phrase-reading task.

The reading task in Brazilian Portuguese intended to screen participants for dialectal and idiolectal variations such as palatalization of /s/ and /z/ and contained nine sentences with words such as ‘esperto’ (smart in English) and ‘diferente’ (different in English) which are words subject to many variations in pronunciation across Brazilian Portuguese dialects. Palatalization of the fricatives was not found in the data collected, though. Participants were from the three states in the south of Brazil.

The short interview aimed at eliciting word-initial /s/-clusters through a less controlled task and, thus, more similar to daily language production than the reading tasks. There were fourteen questions that could trigger answers containing word-initial /s/-clusters. The paragraph-reading task is available on the internet (<http://accent.gmu.edu/pdfs/elicitation.pdf>) and it was designed to screen native and non-native English accents. The word-initial /s/-clusters in the paragraph are ‘Stella’, ‘store’, ‘spoons’, ‘snow’, ‘slabs’, ‘snack’, ‘small’, ‘snake’, ‘scoop’, and ‘station’. The story-reading task is *The story of Sleeping Beauty* and it is also available on the internet (<http://www.bbc.co.uk/education/wordsandFigures>). Since, word-initial /s/-clusters are usually mispronounced by very young native speakers of English, it was designed to provide native English-speaking children with practice on producing such target. Sixteen tokens of word-initial /s/-clusters are included in the story.

The phrase-reading task administered in Bettoni and Koerich (2009) was used to assess production in controlled phonological contexts including phrases used in the perceptual training task as well as untrained phrases different from the story and paragraph reading which varied in terms of number of each cluster. The clusters included were /sl, sm, sn, sk, sp, st, skr, spr/. There were two words representing each cluster included in the training program (/sl, sm, sp, st/) and one word representing each cluster not included in the training program. Four preceding phonological contexts were tested (a) a vocalic context as in ‘how smiles’; (b) a voiceless consonant as in ‘if smiles’; (c) a voiced consonant as in ‘move smiles’; and (d) silence as in ‘smiles’. The total of tokens was 45 per participant.

### **2.3 Perceptual identification training program**

The training task was the same used in Bettoni and Koerich (2009). The stimuli used in the training were recorded by two Americans: a male and a female. The part of the stimuli (four clusters) not used in the training were used to investigate generalization. The inclusion of two talkers aimed at adding variability so that within categories distinctions were reduced.

There were eight pairs of trained contrasts. There were two words for each trained cluster. For each word, there were four preceding contexts: a vowel-like sound as in ‘how’, a voiced consonant as in ‘move’, a voiceless consonant as in ‘if’, and silence. The training material consisted of two-alternative-forced-choice (2AFC) identification trials with immediate feedback and replay allowed before hitting the decision key.

The training program consisted of six sessions with numbers of blocks varying from one to twenty-four. The number of trials in each block varied from sixteen to sixty-four. The number of talkers, number and type of clusters, number and type of contrasts, and number and type of contexts were presented in increasing difficulty in the first three sessions (one talker alone before two talkers, /s/+stop before /s/+sonorant clusters; prothesis before voicing; and, silence before consonant before vowel). The hierarchy of difficulties concerning types of clusters and types of contexts was based on Rebello (1997).

The first session contained stimuli recorded by only one of the talkers. The second session contained stimuli recorded by the other talker. From the third session on, stimuli produced by both talkers were put together and presented randomly. In the fourth session, contexts were randomized for presentation. In the fifth session, contexts, as well as contrasts, were randomized. Finally, in the sixth session, clusters, contrasts, and contexts were randomized.

The participants were allowed to advance to the following block only after having obtained 93% of accuracy in the first three sessions and after 91% of accuracy in the other sessions (adapted from ROCHET, 1995). The difference in the level of accuracy required was due to the different number of trials in the blocks across sessions.

At the end of each session, there was an imitation block featuring some of the twenty-eight phrases containing /s/-clusters used in the training. The imitation block was included with the intent of guiding participants' attention to the relationship between perception and production, since research has shown that attention guiding is an efficient strategy (e.g., GUION; PEDERSON, 2007). Participants listened to a phrase, imitated it, and read on the laptop screen the just-imitated phrase.

## 2.4 Procedure

Even though participants could communicate very well in English, they were instructed in Brazilian Portuguese in order to ensure comprehension. The study consisted of five phases: (a) a pretest; (b) a training phase; (c) a posttest; (d) an eight-month retention test; and (e) a twelve-year retention test.

During data gathering in order to test other hypotheses which are not the focus of the present study, an identification task and a discrimination task were administered in phases a, c, d, and e. For the purposes of the present research article only results and procedures concerning production are discussed. The twenty-three participants took the pretest individually and were given short breaks after each task. They answered the questionnaire, then the production tasks followed by the perception tasks which are not discussed here (for details see BETTONI-TECHIO, 2008). The sessions lasted from 45 to 60 minutes depending on the participant. Following that, the participants were individually trained on the identification of word-initial /s/-clusters. The first block of training consisted of the easiest cluster in the easiest phonological testing the most salient contrast (if stop vs. iffy stop). According to each participant's performance on the first block of training, the participant would be trained on the 'prothesis' contrast. A 100 % of accuracy in the first block of training determined that the participant skipped the blocks where 'prothesis' was the only pertinent contrast; however, he/she had to take the blocks which trained 'prothesis', 'voicing', and 'prothesis and voicing' contrasts altogether. This measure was taken to allow the participants to focus on their specific difficulties, individualizing the training and avoiding exhaustion by unnecessary repetition. At the same time, a minimum amount of exposure to all contrasts was ensured. The total length of training varied from two to six hours.

The twenty-three participants took the posttest around ten days after the pretest, that is, immediately after the training administered to the fifteen participants in the experimental group and with a ten-day gap for the control group. The order of the tests was (1) the interview, (2) the paragraph-reading task, (3) the story-reading task, (4) the phrase-reading task, (5) the AX discrimination task, and (6) the identification task. Participants took from 40 to 50 minutes to complete the posttest.

The first retention test was identical to the posttest and was administered around eight months later. Eight participants from the experimental group (P1, P2, P4, P5, P8,

P9, P10, and P12) took the retention test. The other participants either could not be reached or failed to attend to the retention test. The second retention test had all the tasks in the first retention test and an additional interview about habits and exposure to English during the twelve-year gap between the tests. P1, the youngest participant in the group, was the only participant who could be reached. She was nine at the time of the first data gathering phase and twenty-one at the twelve-year retention test. The administration of all tests and the training program were conducted by the researcher in individual meetings with the participants.

## 2.5 Data analysis

Participants' production was perceptually analyzed by the researcher, and acoustically analyzed using Praat 4.3.12 software for the pretest, the posttest and the eight-month retention test and Praat 6.1.16 for the twelve-year retention test. With the aid of spectrograms, the productions of /sC(C)/ were categorized as /sC(C)/ meaning target-like production, /VsC(C)/ meaning production with prothesis, /zC(C)/ meaning production with voicing or /VzC(C)/ meaning production with prothesis followed by voicing.

The data received statistical treatment using the Statistical Package for Social Studies (SPSS) software in order to better investigate the proposed hypotheses. Wilcoxon Signed Ranks Tests were selected to investigate improvement. Mann-Whitney tests were used for comparing gain scores and proportion of the room for improvement used. Besides gain scores, proportions of the room for improvement were calculated according to one of the following formulae:  $(\text{posttest}-\text{pretest})/(100-\text{pretest})$  – when dealing with percentages, or  $(\text{posttest}-\text{pretest})/(N-\text{pretest})$  – when dealing with raw scores.

## 2.6 Results of the control group

A control group was included in the study in order to set a parameter for the improvement which could be found as the result of a training by the experimental group. The intention was to discard (1) improvement caused by task effect, and (2) learning promoted by test-taking. It would not be considered bad that learners benefited from the testing, but it was necessary to understand how much improvement could be attributed to testing itself and how much could be attributed to the training program. Table 1 displays the results in the pre- and posttests for the control group.

**Table 1: Results in the pre- and posttests for the control group**

	N	Pretest		Posttest	
		Score	Rate	Score	Rate
P16C	71	31	43.66%	40	56.34%
P17C	71	16	22.54%	16	22.54%
P18C	71	17	23.94%	17	23.94%
P19C	71	71	100%	71	100%
P20C	71	32	45.07%	40	56.34%
P21C	71	53	74.65%	55	77.46%
P22C	71	62	87.32%	61	85.92%
P23C	71	25	35.21%	27	38.03%
<b>MEAN</b>		<b>38</b>	<b>54.05%</b>	<b>41</b>	<b>57.57%</b>

The control group performed worse than the experimental group in the pre and posttest, and therefore had a larger room for improvement. According to Yeon (2004),

weaker learners benefit more from training, so participants in the control group had a greater chance of improvement than participants in the experimental group. Contrary to that, however, the gain score for the control group in production was only 3.53%. The proportion of the room for improvement used was 7.66% indicating that 93.33% of room for improvement was not used.

Wilcoxon Signed Ranks tests run on the data yielded non-significant results ( $p > .05$ ),  $Z = 1.761$ . Since four out of eight participants had a slight improvement in production, it can be concluded that these learners may have had some awareness of the error they were producing in the pretest through the tests. However, as the statistical tests indicated, this awareness was not sufficient to form a new sound category. Thus, the improvement found for the experimental group, as will be seen later, can be attributed to training effects.

### 3. ANALYSIS AND DISCUSSION OF THE RESULTS

Results are presented and discussed for each of the four hypotheses in the order they were introduced.

#### 3.1 Improvement in production

In order to verify whether there was improvement in production, results from the four production tasks in English will be discussed: the story-reading task, the paragraph-reading task, the phrase-reading task, and the interview. Also, the overall production results excluding the interview data were discussed since number and type of clusters produced in the interview varied among the participants. Results from pre and posttests for reading tasks are displayed by each participant in the experimental group in Table 2.

**Table 2: Results for the reading tasks in the pre- and posttests for the experimental group**

	N	Pretest		Posttest		Gain Score	P*
		Score	Rate	Score	Rate		
P1	71	64	90.14%	68	95.77%	5.63%	57.14%
P2	71	59	83.10%	71	100%	16.90%	100%
P3	71	40	56.34%	67	94.37%	38.03%	87.10%
P4	71	32	45.07%	65	91.55%	46.48%	84.62%
P5	71	47	66.20%	67	94.37%	28.17%	83.33%
P6	71	34	47.89%	57	80.28%	32.39%	62.16%
P7	71	62	87.32%	67	94.37%	7.05%	55.56%
P8	71	53	74.65%	67	94.37%	19.72%	77.78%
P9	71	48	67.61%	68	95.77%	28.16%	86.96%
P10	71	34	47.89%	65	91.55%	43.66%	83.78%
P11	71	57	80.28%	71	100%	19.72%	100%
P12	71	62	87.32%	71	100%	12.68%	100%
P13	71	71	100%	71	100%	0	0
P14	71	57	80.28%	68	95.77%	15.49%	78.57%
P15	71	46	64.79%	68	94.37%	29.58%	88.00%
<b>MEAN</b>		<b>51</b>	<b>71.92%</b>	<b>67</b>	<b>94.84%</b>	<b>22.92%</b>	<b>80.00%</b>

P\* stands for proportion of the room for improvement used.

The gain score in production was 22.92% which corresponded to 80% of the room for improvement. Besides the perceptual training, participants' production benefited from

the awareness of the target production of word-initial /s/-clusters which provided them with tools to monitor their own speech. The awareness mentioned here refers to the awareness raised by the training program and by metalanguage. Mere awareness of the rule tends not to lead to better pronunciation (Barbara Baptista in personal communication, 2006),

A Wilcoxon Signed Ranks test run on the production scores for the reading tasks of pre- and posttest yielded a significant result,  $Z = 3.297$ ,  $p = .001$ . Comparing this result to the one found for the control group, this corroborates the hypothesis that there would be improvement in production as an effect of the perceptual training. Considering the reading tasks individually, Wilcoxon Signed Ranks Tests yielded significant improvement for the paragraph-reading task:  $Z = 3.201$ ,  $p = .001$  and for the story-reading task:  $Z = 2.810$ ,  $p = .005$ . Eight participants obtained a perfect score in the phrase-reading test in the posttest. A Wilcoxon Signed Ranks test indicated that there was improvement in the phrase-reading task –  $Z = 3.300$ ,  $p = .001$  as well.

By comparing the Z scores obtained for the three tasks, it can be noticed that the phrase-reading task was the task with the highest improvement followed by the paragraph-reading test, and then by the story-reading task. Thus, the tasks in which the possibility of control and monitoring were higher had the greatest improvement.

Considering the type of mispronunciation, Hypothesis 2 suggested production of word-initial /s/+sonorant clusters would remain more problematic than production of word-initial /s/+stop clusters. /s/+sonorant clusters are more prone to voicing and prothesis followed by voicing. The data concerning type of mispronunciation in both pre- and posttests is displayed in Table 3.

**Table 3: Type of mispronunciation of word-initial /s/-clusters in pre and posttests**

	N	Prothesis		Voicing		Prothesis + Voicing		Total
		Score	Rate	Score	Rate	Score	Rate	
Pretest	1065	55	5.16%	178	16.71%	71	6.67%	304
Posttest	1065	14	1.31%	30	2.81%	6	0.56%	50

Wilcoxon Signed Ranks tests showed that the reductions from the pretest to the posttest were all significant –  $Z = 2.823$ ,  $p = .005$ , for prothesis,  $Z = 3.300$ ,  $p = .001$ , for voicing, and  $Z = 2.937$ ,  $p = .003$ , for prothesis combined with voicing. A Wilcoxon Signed Ranks test run on “proportion of the room for improvement” for prothesis and voicing was significant –  $Z = 2.340$ ,  $p = .013$  – showing that the reduction in the occurrence of voicing was significantly larger than the reduction in the occurrence of prothesis. Even though voicing tends to affect only /s/+sonorant clusters and prothesis tends to affect all cluster types, voicing was three times more frequent than prothesis in the pretest, and two times in the posttest. The occurrence of prothesis combined with voicing was also larger than of prothesis alone in the pretest. In order to have a clear view of the /s/+sonorants, which are the ones which typically suffer voicing and prothesis followed by voicing, and address generalization of training, the data was organized in Table 4 considering trained and untrained /s/+sonorant clusters.



**Table 4: Production of /s/+sonorant clusters in the reading tasks**

	N	Pretest		Posttest		Gain Sc.	P*
		Target-like	RATE	Target-like	RATE		
Trained cluster							
/sm/	150	63	42.00%	141	94.00%	52.00%	89.65%
/sl/	165	90	54.54%	156	94.54%	40.00%	88.00%
Untrained cluster							
/sn/	195	98	50.25%	154	78.97%	28.72%	57.77%
<b>Total</b>	<b>510</b>	<b>251</b>	<b>49.21%</b>	<b>451</b>	<b>88.43%</b>	<b>39.22%</b>	<b>77.22%</b>

\*P stands for proportion of the room for improvement

Out of 304 mispronunciations in the pretest, 259 occurred when the target was an /s/+sonorant cluster. For the posttest, 19 out of 50 mispronunciations occurred in /s/+sonorant clusters. For the /s/+stop clusters the reduction in mispronunciations was from 45 to 31. The accuracy rate in the pretest was 91.89% and in the posttest was 94.41%. The proportion of room for improvement used for the /s/+stop clusters was, then, 44.93%. The accuracy rate in the pretest for the /s/+sonorant clusters was 49.21% and in the posttest 88.43%. The proportion of room for improvement used for the /s/+sonorant clusters was, then, 77.22%. Even though the gain score was higher for the /s/+sonorants and more room for improvement was used, the /s/+sonorant clusters remained more problematic for the participants of the research corroborating Hypothesis 2.

Mann-Whitney Tests run on gain scores and on proportion of the room for improvement for trained and untrained clusters in production failed to yield significance,  $Z = .227$ ,  $p = .820$  and  $Z = .602$ ,  $p = .547$  respectively. The results confirmed that there was generalization of improvement in production to the untrained clusters.

Considering the four production tasks administered, the interview was closer to natural speech than the reading tasks. One of the characteristics of natural speech is the variability in type and amount of production. Even though the questions used in the interview were prepared in a way to elicit the production of word-initial /s/-clusters, the answers varied from one participant to another and from the pre- to the posttests because the researcher had no strict control over the productions. The words elicited beginning in /s/+stop were 'school', 'special', 'spent', 'spoon', 'stable', 'Stallone', 'stars', 'starting', 'stay', 'stereo', 'still', 'stop', 'storm', 'stove', 'student', 'study', 'studying', and 'stuff'. The words beginning in /sCC/ elicited were 'spring', 'street', and 'strong' and the words beginning in /s/+sonorant elicited were 'sleep', 'slow', 'slowly', 'small', 'smile', 'smiles', and 'snow'.

Few modifications in /s/+stop clusters occurred – four in the pretest and none in the posttest. The number of errors was small and 100% of the room for improvement was used. The number of modifications in /sCC/ was even smaller than in /s/+stop – there was only one error in the pretest. However, the number of productions obtained was smaller as well. Again, 100% of the room for improvement was used. As in the reading tasks, /s/+sonorant clusters were more frequently modified in the interview. Since the rates in the pretest were low, the room for improvement was large and so were the gain scores obtained from the pre- to the posttest - /sm/: 56%; /sn/: 20%; and /sl/: 55%. The data obtained is organized by cluster and type of mispronunciation for both pretest and posttest in Table 5.

**Table 5: Mispronunciations of word-initial /s/-clusters in the interview in the pre and posttests**

	Pretest					Posttest				
	N	z	Vs	Vz	%-error	N	z	Vs	Vz	%-error
sp*	11	0	0	0	-	5	0	0	0	-
st*	81	0	3	1	<b>4.94%</b>	76	0	0	0	-
sk	1	0	0	0	-	0	0	0	0	-
spr	21	0	0	0	-	19	0	0	0	-
str	15	0	1	0	<b>0.67%</b>	15	0	0	0	-
sm*	20	10	1	4	<b>75.00%</b>	21	4	0	0	<b>16.67%</b>
sn	15	3	0	2	<b>33.33%</b>	16	2	0	0	<b>12.50%</b>
sl*	19	12	1	1	<b>73.68%</b>	21	3	0	1	<b>14.28%</b>
<b>TOTAL</b>	<b>184</b>	<b>25</b>	<b>6</b>	<b>8</b>	<b>21.19%</b>	<b>173</b>	<b>9</b>	<b>0</b>	<b>1</b>	<b>5.78%</b>

z = voicing of word-initial /s/-cluster

Vs = prothesis of word-initial /s/-cluster

Vz = prothesis of word-initial /s/-cluster followed by voicing

\* = trained clusters

The discussion about the influence of the level of formality of speech in L2 pronunciation accuracy is far from being resolved. Investigating final epenthesis production with Brazilian learners of English, Major (1986) and Koerich (2002) found that errors tended to decrease as the formality of the task increased (from the reading of texts, to sentences, and then to lists of words, in the former study and from free-speech to sentence-reading in the latter). Statistical treatment of Koerich's data showed that the difference between rates of epenthesis were not significant though.

In the present study, the data revealed that, for some participants, the number of errors in the interview was smaller than in the other tasks – the paragraph-reading task and the phrase-reading task. Although the issue was not deeply studied, it might be reasoned that more mispronunciations presented in the more formal tasks (contrary to the two studies mentioned above) resulted from the combinations of difficult contexts and clusters in the texts, which did not occur in the interview.

The participants of the present study, especially those in the experimental group, had high level of language proficiency and presented a higher rate of 'voicing' than of 'prothesis'. Thus, the data might be taken to indicate that voicing is a more resistant error, whereas prothesis and paragoge (initial and final epenthesis) tend to be more common in the pronunciation of learners with lower proficiency levels (MAJOR, 1986, 1992, 1996; KOERICH, 2002).

### 3.2 Retention of improvement after eight months

Studies have investigated retention with one, three, five, six month and even three-year follow-up tests. Studies investigating six-month retention usually have a three-month retention test as well. This measure might change results since learners are called attention to the contrast in the middle of the six-month period. The eight-month interval of the present study was challenging. One of the greatest difficulties of carrying out such a long-term study is the availability of participants. As explained in the method, only eight out of fifteen participants took the eight-month retention test. The data gathered is displayed in Table 6.

**Table 6: Results for the reading tasks in the pre- and posttests for the experimental group**

	N	Pretest		Posttest		Eight-month		Eight - Pretest	
		Score	Rate	Score	Rate	Score	Rate	Gain Sc	P*
P1	71	64	90.14%	68	95.77%	71	100%	9.76%	100%
P2	71	59	83.10%	71	100%	65	91.55%	8.45%	50%
P4	71	32	45.07%	65	91.55%	58	81.69%	36.62%	66.67%
P5	71	47	66.20%	67	94.37%	63	88.73%	22.53%	66.67%
P8	71	53	74.65%	67	94.37%	68	95.77%	21.12%	83.31%
P9	71	48	67.61%	68	95.77%	59	83.10%	15.49%	47.82%
P10	71	34	47.89%	65	91.55%	67	94.37%	46.48%	89.19%
P12	71	62	87.32%	71	100%	71	100%	12.68%	100%
<b>MEAN</b>			<b>70.25%</b>		<b>95.38%</b>		<b>91.90%</b>	<b>21.65%</b>	<b>72.78%</b>

Wilcoxon Signed Ranks tests run on the pre- and retention tests –  $Z = 2.521$ ,  $p = .012$  – and on the post- and retention tests –  $Z = 1.352$ ,  $p = .176$  – indicated that the improvement was retained after eight months corroborating the third hypothesis. Three participants remained improving between posttest and the eight-month follow-up test. One explanation may be that perceptual training triggered production monitoring strategies which allowed participants to remain developing their phonological representations of the target. One participant who had a perfect score in the posttest remained with the perfect score and the other four participants had a slight reduction in their scores comparing post and retention test, but all of them performed better in the retention test than in the pretest.

Considering the interview, P1 voiced /sm/ in the posttest and showed improvement in the retention test by achieving 100% accuracy. P5 also voiced /sm/ in the posttest and did not voice /sm/ in the retention test. P5 had not produced /sl/ in the pre- and posttests, but produced two tokens of it in the retention test – one was accurate and the other was voiced. P9 produced voicing in combination with prothesis in /sl/, and voicing in /sm/ in the posttest and only voicing in /sl/ in the retention test. P2, P4, P8, P10, and P12 maintained 100% accuracy shown in the posttest. Therefore, all participants who had some room for improvement, improved from the posttest to the retention test, and the others maintained the perfect score.

### 3.3 Retention of improvement after twelve years

As previously mentioned, the only participant who could be reached for the twelve-year retention test was P1. The, then, nine-year-old child had been exposed to English in and outside class for a few years and presented the same types of mispronunciations triggered by Brazilian Portuguese phonotactics of the other participants of the research at the pretest phase. Considering the time lapse, an informal interview was carried out to collect information about P1's experience with English along the years. The data gathered in the interview concerned the following topics: formal English instruction, experience in an English speaking country, and English language use – perception and production.

Regarding formal English instruction, P1 stopped attending English language classes at extracurricular courses one year after the pretest. She, then, started having French classes. She is fluent in French nowadays. She attended English classes in regular school until she was seventeen and was admitted into college. Six years after dropping

the English course, when she was sixteen years old, she attended four months of a Cambridge Advanced Exam preparation course, took the test, and got an A. When P1 was fourteen years old, she visited the U.S.A. and stayed there for two weeks. She stated that she interacted mostly with her Brazilian peers and in Portuguese.

P1's exposure to spoken English is high. Her favorite bands and singers sing in English. She watches sitcoms and drama series in English very often. She never watches dubbed audiovisual productions and sometimes make use of subtitles even though she can understand spoken English very well. She is also exposed to Spanish, French, and Portuguese audiovisuals, but most of the material she watches is in English. She has few opportunities for oral interactions in English and, thus, does not speak English frequently.

After the interview, which was carried out in Brazilian Portuguese, P1 recorded the paragraph reading task, the Sleeping Beauty story reading task, the phrase reading task and was interviewed in English being asked the same questions of the interviews carried out in the pretest, posttest and eight-month retention test which aimed at assessing word-initial /s/-clusters production in a less controlled situation.

The 71 instances of word-initial /s/-clusters in the reading tests were produced in a native-like fashion according to both perceptual and acoustic analysis. Twenty-one instances of word-initial /s/-clusters were produced in the twelve-year retention interview, almost twice the number of instances of the eight-month retention interview which was eleven. All instances were accurately produced confirming the fourth hypothesis. The data comparing pre, post and retention tests according to type of mispronunciation produced by P1 across time is displayed in Table 7.

**Table 7: P1 production of word-initial /s/-clusters in the four testing phases**

<b>Reading tasks</b>		<b>Target-like</b>		<b>Voicing /zC/</b>		<b>Prothesis /VsC/</b>		<b>/VzC/</b>	
	N	Score	Rate	Score	Rate	Score	Rate	Score	Rate
Pretest	71	64	90.14%	5	7.04%	2	2.81%	0	0
Posttest	71	68	95.77%	3	4.22%	0	0	0	0
8-month	71	71	100%	0	0	0	0	0	0
12-year	71	71	100%	0	0	0	0	0	0
<b>Interview</b>		<b>Target-like</b>		<b>Voicing /zC/</b>		<b>Prothesis /VsC/</b>		<b>/VzC/</b>	
	N	Score	Rate	Score	Rate	Score	Rate	Score	Rate
Pretest	18	13	72.22%	4	22.22%	0	0	1%	5.55%
Posttest	10	9	90%	1	10%	0	0	0	0
8-month	11	11	100%	0	0	0	0	0	0
12-year	21	21	100%	0	0	0	0	0	0

Besides the effortless and accurate word-initial /s/-clusters production, analysis of all data produced by P1 raises a concern. Even though P1's pronunciation in English was intelligible, accurate and fluent, her production of final 's' when marking third person singular and plural forms did not follow the voicing rules of English. She effortlessly produces the /s/ in word-initial /s/-clusters without voicing, palatalization or epenthesis. However, she fails to voice final grapheme 's' when it should be voiced as in the end of the words 'peas' and 'bags'. One hypothesis is that there was an overgeneralization for the letter 's' in word initial /s/-clusters being always produced as /s/. Another hypothesis is that she simply failed to acquire naturally the voicing pattern of final 's' in English and has not been made aware of the differences between Brazilian Portuguese and English phonological rules concerning voicing and final /s/ production.

#### 4. FINAL REMARKS

The main objective of the present study was to investigate effects of perceptual training on the pronunciation of word-initial /s/-clusters in the Brazilian Portuguese/English interphonology. The first hypothesis that there would be improvement in production from the pretest to the posttest administered immediately after the perceptual training was confirmed since the gain score from pre to posttest was 22.92% (from 71.92% to 94.84%) which corresponded to 80% of the room for improvement. Participants improved differently in the different tasks. Since participants had a minimum B1 level of proficiency in English, prothesis was not so frequent. Also, palatalization was not produced. This may be explained by the results in the Brazilian Portuguese reading task which showed that palatalization was not present in their idiolects in Brazilian Portuguese. As Bettoni-Techio (2005) stated, not everyone who palatalizes in Portuguese transfers this phonological process to English, but the ones who use palatalization in English where it is not the target pronunciation, have palatalization in their L1 idiolect.

Voicing, on the other hand, proved to be a persistent mispronunciation. Regarding the constituents of the clusters, /s/+sonorant clusters were significantly more difficult than /s/+stop clusters corroborating Cornelian (2003), Rauber (2002, 2006), and Rebello and Baptista (2006). In the pretest, 85.19% (259) of the mispronunciation occurred when the target was word-initial /s/+sonorant clusters and voicing was the commonest mispronunciation. Nearly all participants had a close to perfect score on the phrase-reading task in the posttest where more control was allowed, and monitoring strategies took place. For the interview, in the posttest, there were no mispronunciations of /s/+stop clusters and there was a considerable reduction of 46.19% in the number of mispronunciations of /s/+sonorant clusters. Voicing, however, was still present in 14.48% of the /s/+sonorant productions. The number of mispronunciations was significantly reduced from pre- to posttest, but voicing of /s/+sonorants remained more persistent for the participants of the research corroborating Hypothesis 2.

Retention of improvement in the production domain was tested twice. The first time was eight months after the training program. Eight out of the fifteen original participants of the experimental group participated, and the gain score obtained from pre to retention test was 21.65% for the reading tasks. Wilcoxon Signed Ranks tests run on the pre- and retention tests –  $Z = 2.521$ ,  $p = .012$  – and on the post- and retention tests –  $Z = 1.352$ ,  $p = .176$  – indicated that the improvement was retained after eight months corroborating the third hypothesis. For the twelve-year retention test, only one participant of the perceptual training program could be reached and even though she obtained the same perfect score she had obtained in the eight-month retention test, she produced nearly the double of the word-initial /s/-clusters productions in the interview. All of her productions of word-initial /s/-clusters were target-like confirming the fourth hypothesis. The participant stated that she had had massive exposure to spoken English throughout the years. Even though, her success in the twelve-year retention test could be attributed to that, other mispronunciations which are not as persistent as voicing of word-initial /s/-clusters are present in her speech, such as non-target-like production of final ‘s’ in plural forms. This may suggest that, for non-native speakers, mere exposure to spoken English is not sufficient and awareness raising may be necessary for many people. It serves as a reminder that people do not acquire a second language the same way they acquire their first one. Our first and second languages interfere in each other.

Overall, perceptual training of /s/-clusters caused positive changes in production, since there was improvement in the reading tasks and in the interview for trained and

untrained clusters. Results of posttest and of the retention tests may serve as evidence that adults are able to learn new sounds through massive exposure to the target language and immediate feedback – the strategies used in the training program. Participants started relying on cues that were ignored before training. The overall results indicate that there may be a common mechanism for perception and production, “since both domains are affected by the same variables and to a similar extent” (BETTONI; GALLEGO-CAMPOS, 2015).

---

## REFERENCES

- ALVES, U. K.; KAMPPFF, F. R. Efeitos de longo prazo do treinamento perceptual na percepção e produção das plosivas iniciais surdas do inglês por estudantes brasileiros: implicações para o ensino da pronúncia. *Ilha do Desterro* v. 72, nº 3, p. 375-399, Florianópolis, set/dez 2019. doi:10.5007/2175-8026.2019v72n3p375
- BETTONI, M.; GALLEGO-CAMPOS, F. R. Preceding phonological context and cluster type effects on word-initial /s/-cluster pronunciation in Brazilian Portuguese/English interphonology. *Revista Letras*, Vol. 92, p. 133-152, 2015.
- BETTONI, M.; KOERICH, R. D. Perceptual training on the pronunciation of /s/-clusters in Brazilian Portuguese/English In: RAUBER, AS, WATKINS, MA; BAPTISTA, BO. *Recent Research in Second Language*, 154-173., 2009.
- BETTONI-TECHIO, M. *Production of final alveolar stops in Brazilian Portuguese/English interphonology*. Unpublished Master Thesis. Federal University of Santa Catarina, 2005.
- BETTONI-TECHIO, M. *Perceptual training and word initial /s/-clusters in Brazilian Portuguese/English interphonology*. Unpublished doctoral dissertation. Federal University of Santa Catarina, 2008.
- BRADLOW, A. R.; PISONI, D. B.; AKAHANE-YAMADA, R.; TOHKURA, Y. Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America* 101 (4), pp. (2299 – 2310), 1997.
- COOKE, M.; GARCIA LECUMBERRI, M. L. Effects of exposure to noise during perceptual training of non-native language sounds. *The Journal of the Acoustical Society of America*, 143(5), 2602–2610, 2018. doi:10.1121/1.5035080
- CORNELIAN, D. *Brazilian Learners’ Production of Initial /s/ Clusters: Phonological Structure and Environment*. Unpublished master thesis. Federal University of Santa Catarina, 2003.
- FLEGE, J. E. Chinese subjects’ perception of the word-final English /t/-/d/ contrast: performance before and after training. *The Journal of the Acoustical Society of America*, 86, (5), 1984-1697, 1989.
- GUION, S; PEDERSON, E. Investigating the role of attention in phonetic learning. In O.-S. Bohn & M. Munro (Eds.) *Second-language speech learning: The role of language experience in speech perception and production: A festschrift in honour of James Emil Flege*. (pp. 57-77) Amsterdam: John Benjamins., 2007.
- HARDISON, D. Transfer of second-language perceptual training to production improvement: focus on /r/ and /l/. In: James, A. & Leather, J., eds., *New Sounds 2000: Proceedings of the Fourth International Symposium on the Acquisition of Second-Language Speech*. University of Klagenfurt, 166-173, 2000.
- HARDISON, D. Acquisition of second-language speech: Effects of visual cues, context, and talker variability. *Applied Psycholinguistics*, 24, 495-522, 2003.
- JAMIESON, D. G.; MOROSAN, D. E. Training nonnative speech contrasts in adults: Acquisition of the English /ð/ -/θ/ contrast by francophones. *Perception and Psychophysics*, 40 (4), 205-215, 1986.
- KABAKOFF, H; GO, G.; LEVI, S. V. Training a non-native vowel contrast with a distributional learning paradigm result in improved perception and production. *Journal of Phonetics*, vol. 78, 100940, 2020. <https://doi.org/10.1016/j.wocn.2019.100940>
- KOERICH, R. D. *Perception and production of word-final vowel epenthesis by Brazilian EFL students*. Unpublished doctoral dissertation. Universidade Federal de Santa Catarina, 2002.
- LAMBACHER, S. G.; MARTENS, W. L.; KAKEHI, K.; MARASINGHE, C. A.; MOLHOLT, G. The effects of identification training on the identification and production of American English vowels by native speakers of Japanese. *Applied Psycholinguistics*, 26, 227-247, 2005.

- LIVELY, S. E.; LOGAN, J. S.; PISONI, D. S. Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America*, 94, 1242-1255, 1993.
- MAJOR, R.C. The natural phonology of second language acquisition. In A. James & J. Leather (Eds.) *Sounds Patterns in Second Language Acquisition* (pp. 207-224). Dordrecht: Foris, 1986.
- MAJOR, R. C. Transfer and developmental factors in second language acquisition of consonant clusters. Paper presented at the New Sounds 92: Proceedings of the 1992 Amsterdam symposium on the acquisition of second-language speech, Amsterdam, Holland, 1992.
- MAJOR, R.C. Markedness in second language acquisition of consonant clusters. *Studies in Bilingualism*, 10, 75-96, 1996.
- RAUBER, A. S. *The Production of English Initial /s/ Clusters by Portuguese and Spanish EFL Speakers*. Unpublished master thesis. Universidade Federal de Santa Catarina: Florianópolis, 2002.
- RAUBER, A. S. Production of English initial /s/-clusters by speakers of Brazilian Portuguese and Argentine Spanish. In B. O. Baptista & M. A. Watkins. *English with a Latin Beat: Studies in Portuguese/Spanish – English Interphonology* (pp. 155-167). Amsterdam: John Benjamins, 2006.
- REBELLO, J. T.; BAPTISTA, B. O. The influence of voicing on the production initial /s/-clusters by Brazilian learners. In B. O. Baptista & M. A. Watkins. *English with a Latin Beat: Studies in Portuguese/Spanish – English Interphonology* (pp. 139-154). Amsterdam: John Benjamins, 2006.
- REBELLO, J. T. The acquisition of initial /s/ clusters by Brazilian EFL learners. In J. Leather & A. James (Eds.) *New Sounds 97: Proceedings of the third international symposium on the acquisition of second-language speech* (pp. 336-350). Klagenfurt: University of Klagenfurt, 1997.
- ROCHET, B. L. Perception and production of second-language speech sounds by adults. In W. Strange (Ed.) *Speech perception and linguistic experience: Issues in cross-language research*. Timonium: York Press, 379-410, 1995.
- STRANGE, W.; DITTMAN, S. Effects of discrimination training on the perception of /r-l/ by Japanese adults learning English. *Perception & Psychophysics*, 36 (2), 131-145, 1984.
- TRAPP, N. L.; BOHN, O. S. Training Danish listeners to identify English word-final /s/ and /z/: Generalization of training and its effect on production accuracy. In: James, A. & Leather, J., eds., *New Sounds 2000: Proceedings of the Fourth International Symposium on the Acquisition of Second-Language Speech*. University of Klagenfurt, 343-350, 2000.
- YEON, S. *Teaching English word-final alveolopalatal to native speakers of Korean*. Unpublished doctoral dissertation. University of Florida, 2004.
- YILDIZ, Y. The structure of initial /s/-clusters: evidence from L1 and L2 acquisition. In Marina Tzakosta, Claartje Levelt and Jeorien van der Weijer (eds.). *Developmental Paths in Phonological Acquisition*. Special issue of Leiden Papers in Linguistics, pp. 163-187, 2005.

Recebido: 4/1/2022

Aceito: 13/10/2022

Publicado: 20/10/2022