An Adaptation of the CDT and Working Memory Capacity

Melissa Bettoni-Techio[1], Denise Cristina Kluge[1], Mara Silvia Reis[1], Denize Nobre-Oliveira[2]

[1]Federal University of Santa Catarina, Brazil
[2]Federal Center of Technological Education of Santa Catarina, Brazil
melissa@formatto.com.br, klugedenise@hotmail.com,
marasreis@hotmail.com, denizenobre@yahoo.com.br

1. Introduction

The instruments used for data gathering are of great importance and should be carefully designed and tested for their validity and reliability. Bachman and Palmer (1996:160-161) state that the investigation of reliability of a language test aims at finding "how much of an individual's test performance is due to measurement error, or to factors other than the language ability we want to measure". Reliability and validity are to be seen as complementary since reliability is a necessary condition for validity (Bachman & Palmer, 1996). However, Bachman and Palmer claim that in order to validate a test one must be concerned with the relationship between test performance and extraneous variables.

Language tests are especially affected by individual differences. Working memory (WM), for instance, is claimed to be a likely component of language aptitude (Miyake & Friedman, 1998). In Brazil, a growing but still limited body of research has been conducted on the relationship between WM capacity and either (a) reading comprehension (e.g., Tomitch, 1995; Torres, 2003; Fontanini, Weissheimer, Bergsleithner, Perucci, & D'Ely, 2005) or (b) speech production (e.g., Fortkamp, 2000; Fontanini et. al, 2005). Research on the relationship between WM capacity and speech perception in Brazilian Portuguese/English interphonology, on the other hand, is scarce (e.g., Rauber, 2004).

Since acquisition of foreign sounds requires a great amount of attention (Leather & James, 1991; Guion & Pederson, 2007) and it is claimed that there is a close relationship between attention and WM (de Fockert, Rees, Frith, & Lavie, 2001), it is reasonable to assume that differences in WM capacity may affect accurate perception of non-native speech contrasts. Also, Grossberg (2003:423) proposes a neural model of speech perception where he claims to exist "top-down attentional interactions between a working memory for short-term storage of phonetic items and a list categorization network for grouping sequences of items".

In order to assess speech perception, researchers have used discrimination and identification tasks (Koerich, 2002; Bettoni-Techio, & Koerich, 2004; Kluge, Rauber, Reis, & Bion, 2007). Flege, Munro, and Fox (1994) designed a categorial discrimination test (CDT) to investigate discrimination of American vowels by non-native speakers. Adaptations of the CDT have been used in studies on Brazilian Portuguese/English interphonology (e.g., Koerich, 2002; Kluge et al, 2007; Silveira, 2004; Bettoni-Techio & Koerich, 2004). In order to give authenticity to the test, instead of using isolated phonemes, these researchers have used words, phrases, or sentences. Thus, an extra load has been imposed on memory and it is unknown whether or not this load compromises the CDT's results. By assessing the relationship between the scores of (a) an adaptation of the CDT, (b) an identification tests, and (c) a span task, reliability and validity of the CDT can be assessed. If WMC is related to perception in general, it will correlate with both perception tests. On the other hand, if WMC correlates with the CDT but not with the identification test, this may indicate that WMC is affecting validity and reliability of this adaptation of the CDT.

The studies previously mentioned (Koerich, 2002; Kluge et al., 2007; Silveira, 2004; Bettoni-Techio & Koerich, 2004) have been carried out mainly on word-final consonants

because of the greater complexity of English syllables in comparison to Brazilian Portuguese (BP) syllables. These studies were mainly concerned with the relationship between perception and production of consonants in coda position. For instance, Koerich (2002) and Silveira (2004) found a correlation between perception and production of word-final consonants using the CDT; Bettoni-Techio, Rauber and Koerich (2007) found a positive and significant correlation between perception and production of word-final alveolar stops also using the CDT; and finally, Kluge et al. (2007) found a significant correlation between perception and production of word-final nasals using both an adaptation of the CDT and an identification test.

Due to differences in syllabic structure and phonetic values between BP and English, perception of word-final nasals is particularly difficult for Brazilian learners. Kluge et al. (2007) found that pre-intermediate learners could neither successfully identify word-final nasals nor discriminate between alveolar and bilabial nasals (see Kluge, 2004 for more details). English native speakers fully pronounce the nasal consonants in syllable-final position by lowering the soft palate and blocking the mouth whereas BP native speakers only lower the soft palate and do not block the mouth producing a nasal vowel instead of a vowel followed by a nasal consonant. This non-native production seems to correlate with inaccurate perception.

In view of the studies and assumptions just reviewed, the present study aims at investigating whether the adaptations of the CDT using words, particularly Kluge et al.'s (2007) adaptation for word-final nasals, are efficient measures of perception or are influenced by working memory capacity.

2. Working memory

According to Baddeley and Logie (1999:28), WM is a multicomponent system made up of a central executive which controls two main slave systems – the phonological loop and the visuospatial sketchpad. Besides its supervisory role, the central executive is also responsible for controlling attention[1] and "activating representations within long-term memory". It is in the phonological loop that speech-based information is manipulated and temporarily stored. And the visuospatial sketchpad is a visual store responsible for the manipulation of actions and images.

Miyake and Shaw (1999:445-449) edited a book on models of working memory containing ten distinct models (Baddeley & Logie; Cowan; Engle, Kane, & Tuholski; Lovett, Reder, & Lebiere; Kieras, Meyes, Mueller, & Seymour; Young & Lewis; Ericsson & Delaney; Barnard; Schneider; O'Reilly, Braver, & Cohen). The editors pointed out six important points of view regarding WM which were shared among all ten models; the points are briefly reported here: (1) there is not a special separate place in the brain for working memory; (2) "WM's maintenance function is in the service of complex cognition" (p. 445); (3) WM is responsible for controlling cognitive actions; (4) multiple factors can affect WM capacity; (5) WM is not a unitary system; and, finally (6) long-term memory and knowledge "play an integral role in WM performance" (p. 449).

Since WM is viewed as a multiple system in service of complex cognition and responsible for controlling attention and temporarily storing information, it can be assumed that individual differences in WM capacity may be a component of language aptitude. In order to test this assumption, WM span tasks were designed. The commonest types of WM span tasks determine the number of items that can be recalled in order following a written

[1] The authors state that the central executive is not the only responsible for controlling attention. They claim that it may "account for some aspects of attention switching and of dividing attention" (p.31).

presentation (e.g., reading span task) or an aural presentation (e.g., listening span task) in which judgment of grammaticality of each sentence is required to assess capacity for controlled attention.

The first reading span task was designed by Daneman and Carpenter (1980). It aimed at verifying whether WM capacity would be a good predictor of reading comprehension performance in L1. The task was designed to test both processing of cognitive actions and storage of information. Even though the first version of the reading span task had its flaws, the task is widely recognized and good adaptations were made to be used with L2 learners (e.g., Harrington & Sawyer, 1992) and via aural presentation (e.g., Mackey, Philp, Egi, Fujii, & Tatsumi, 2002).

In Brazil, studies have been carried out on effects of WM capacity mainly on speaking skills (e.g., Fortkamp, 2000) and reading skills (e.g., Tomitch, 1995; Torres, 2003). Results from these studies are consistent with the claim that WM is a likely component of language aptitude (Harrington & Sawyer, 1992; Miyake & Friedman, 1998).

Based on the studies reviewed, the main objective of the present study is to investigate whether the adaptations of the CDT using words, particularly Kluge et al.'s (2007) adaptation for word-final nasals, are efficient measures of perception or are influenced by working memory capacity. The hypothesis is that adaptations of the CDT using words, for instance, are affected by overload on working memory.

3. Method

3.1 Participants

Participants were 16 master's and doctoral candidates at the Graduate Program in English at UFSC, 15 female and 1 male, aged between 26 and 47 years old, who had already been tested for their proficiency in English in the entrance exam and they were considered advanced speakers of English. The results of the Brazilian advanced learners of English of this study were compared to the results of 20 Brazilian pre-intermediate learners of English (13 women and 7 men, aged between 16 and 44 years), and 3 native speakers of American English (2 women and 1 man, aged between 21-40 years) from Kluge et al.'s study (2007).

3.2 Instruments

For the purpose of data collection, four instruments were administered: two listening span tasks (one in English and one in BP; and two perception tests previously used in Kluge et al., 2007 - an adaptation of the CDT and a native versus non-native identification test).

3.2.1 Listening span tests

Participants' WMC was assessed through a listening span task in English, adapted from Harrington and Sawyer (1992), and a listening span task in BP, adapted from Torres (2003). Both tests consisted of 42 sentences aurally presented in 12 sets: (a) 3 (three) sets of 2 (two) sentences; (b) 3 (three) sets of 3 (three) sentences; (c) 3 (three) sets of 4 (four) sentences; and (d) 3 (three) sets of 5 (five) sentences. The participants had to listen to the sets of sentences and say, right after each sentence, whether it was in a grammatical order or not. After they listened to the whole set of sentences, the participants had to say the last word of each sentence in the order they had listened.

The sentences used in the tests were the same sentences used in Torres for a reading span test (RST) – the author borrowed Harrington and Sawyer's sentences for the RST in English and took the sentences for the RST in BP from popular magazines. Torres' sentences were controlled for number of words, length of final words (nouns – maximum of two syllables), and adapted for ungrammaticality. The addition of a grammatical judgment task to the span test certifies that subjects process the sentences for meaning. The decision of using the same sentences and procedures for the LST was based on Mackey, Philp, Egi, Fujii and Tatsumi (2002), who administered a LST similar to the one used here except for not using sets of two sentences (total: 36 sentences) and for the first language of the participants being Japanese. The ungrammatical sentences corresponded to half of the total sentences and were made so through altering the order of the words which were immediately before the final word.

The sentences were recorded and edited on a laptop computer with the aid of Sound Forge 7.0. The sentences for the LST in BP and in English were recorded by a native BP speaker and a Native American English speaker, respectively. The test was conducted with a Power Point presentation and was semi-randomized. The final words were controlled in order not to have semantically or phonologically related words within a set. Nine extra sentences were recorded by each talker and used as a familiarization task for both tests (LST in BP and LST in English).

3.2.2 Speech perception tests

In order to assess perception, two types of perception tests were used: (a) an adaptation of the CDT used in Kluge et al. (2007), and (b) a native versus non-native identification test used in Kluge et al. (2007).

The CDT consisted of an oddity test format to check if the participants could discriminate final /m/ from /n/. The test was an adaptation of the Categorial Discrimination Test (CDT) designed by Flege, Munro and Fox (2004) to assess ability to perceive English vowels categorically. Participants listened to a total of 72 sets of monosyllabic words. There were 32 "distractor" trials where the target sounds were absent and 40 trials (eight trials for each of the five minimal pairs) with the target nasals that were analyzed in the present study. The trials consisted of three monosyllabic words contrasting either /m/ or /n/ in syllable-final position. In each target set, there were one, two, three, or zero tokens where either /m/ or /n/ was pronounced. Participants had to indicate the odd item out in each trial by circling "1", "2", "3" or "0" if they heard no difference. The audio-stimuli were recorded by three American native speakers of English (one woman and two men).

The native versus non-native identification test consisted of sixty-eight trials of two productions of the same monosyllabic word. The target words were five minimal pairs contrasting final /m/ or /n/ with nasal vowels (non-native pronunciation). The non-native-like English pronunciation consisted of the Brazilian Portuguese nasalization of the vowel and deletion of the consonant. The total of sixty-eight trials consisted of twenty different trials, twenty catch trials (where both words were either native-like or non-native-like) and twenty-eight distractor trials, which were not analyzed in this study. Participants had to indicate which pronunciation sounded more English native-like in each trial by circling "1", "2"; "both" if they considered both pronunciations native-like; or "neither" if they considered neither pronunciation native-like. The participants received an answer grid for the test section with the written word, in order to know what word was being pronounced. The audio-stimuli were recorded by two speakers: one American native speaker of English proficient in BP, and one native speaker of BP proficient in English. Both speakers had phonetic training and could

control their pronunciation so that the nasal was the only difference in the pronunciation of the target words. The two speakers were recorded in individual sessions.

For both tests, speakers were recorded in individual sessions with the aid of Sound Forge 7.0. After normalization for peak intensity, the material was sequenced and edited in Praat software with an inter-trial interval set at 2.8 s and the inter-stimulus at 1.3 s, following Flege (1994). A training session of 12 trials was also designed for the CDT and a training session of eight trials was designed for the identification test.

3.3 Procedure

Data was collected individually on a laptop computer. The first test administered was the CDT and the second one was the identification test because there were written words in the answer sheet and therefore it could compromise the results from the CDT. Then, the LSTs were administered. Half of the participants took the LST in BP first and then the LST in English and the other half did the reverse to avoid ordering effects and consequently a new variable.

Before each of the four tests, participants received instructions and performed a familiarization task. When necessary, the familiarization task was repeated until the participant felt comfortable to take the test. The battery of tests took around 40 minutes. Instructions were given in English for both perception tests and for the LST in English, and in Portuguese for the LST in BP.

3.4 Analysis

The score for the CDT was the number of accurately discriminated target trials, and for the identification test, it was the number of accurately identified target trials. The scores were based on the total of 40 trials for each participant for each of the perception test. For the LSTs, one strict scoring technique was used: the score was the total number of sets where all words were remembered in order (0-12) – All or nothing unit scoring (ALNU) (see Conway, Kane, Bunting, Hambrick, Wilhelm, & Engle, 2005 for details).

Statistical tests were run using the SPSS – version 11.0. Skewness and kurtosis were verified for normal distribution and all results were found to be parametric; thus, the test used to test correlation was the Pearson Product moment of correlation and for comparing means was the Paired sample *t*-test.

4. Results

This study was motivated by the observation that adaptations of the CDT might overload working memory and this may hinder the actual goal of the tests, that is, testing perception of non-native contrasts. The hypothesis was that adaptations of the CDT using words are affected by overload on working memory. In order to test this hypothesis one strict scoring technique for WMC was adopted in the present study (as explained in Section 4.4): the All or Nothing Unit Scoring. A strict scoring technique was adopted because order and form are extremely important for the CDT. Table 1 shows the individual results from the CDT (out of 40 answers) and the identification test (out of 40 answers) as well as the participants' WMC in both Listening Span Tests (out of 12 sets for each test).

Table 1. Individual results for both perceptions tests and the scores of both Listening Span Tests

Part	ID	CDT	ALNUBP	ALNUE
1	20	11	5	5
2	19	20	7	5
3	13	13	4	1
4	17	12	8	7
5	15	32	6	6
6	17	20	9	7
7	23	18	4	6
8	23	21	11	4
9	25	24	11	7
10	31	18	12	5
11	26	20	8	3
12	25	23	6	2
13	22	25	5	2
14	28	26	6	6
15	27	31	4	2
16	31	27	6	4

Acronyms: ID: identification test; CDT: Categorial Discrimination Test; ALNUBP: WMC in BP scored with all or nothing unit score; ALNUE: WMC in English scored with all or nothing unit score.

The results from the CDT and the identification test were correlated with the WM spans and checked for normal distribution through the measurement of skewness and kurtosis. Since the data was normally distributed, as previously mentioned, Pearson correlations were run and yielded no significant results at the .05 level. Figure 1 shows the performance of the advanced learners of English on both perception tests and on both listening span tests.

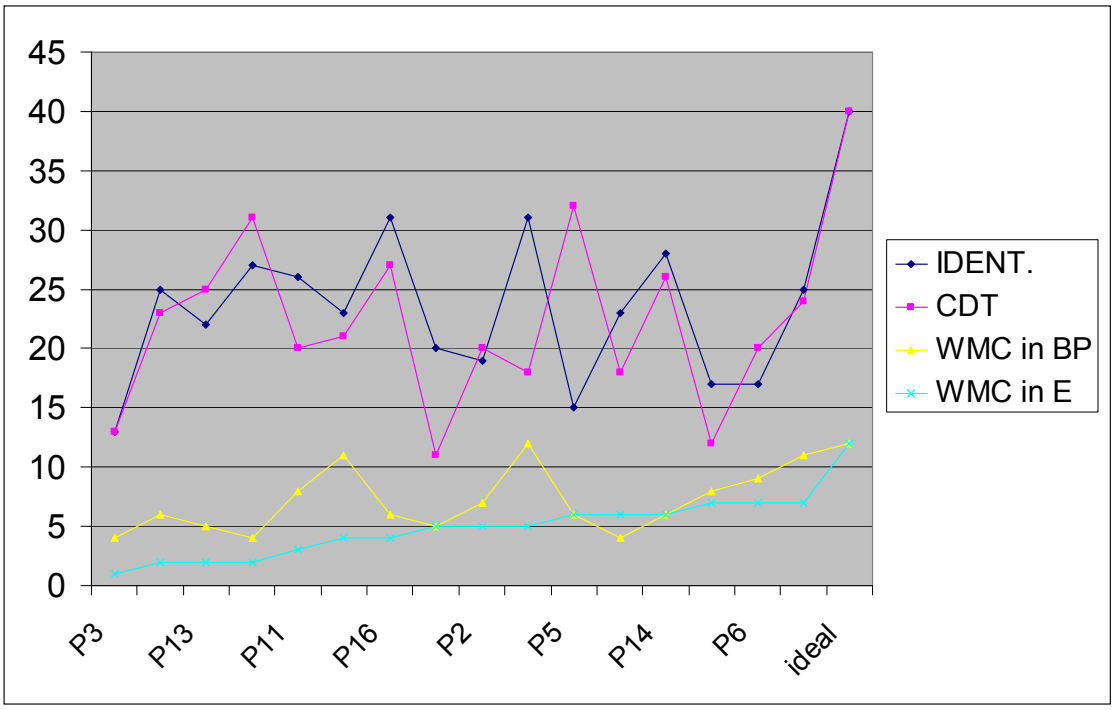

Figure 1. Performance on perception tests and on the listening span tests by participant

Validity was also analyzed by comparing results from a group of native speakers on both perception tests. Also, results of the present study were compared with Kluge et al.'s pre-intermediate learners and native speakers. Therefore, Figure 2 shows the performance of the advanced learners of English of the present study in comparison to the performance of pre-intermediate learners of English and a group of American native speakers of English on the CDT and on the native-like versus non-native-like identification test.

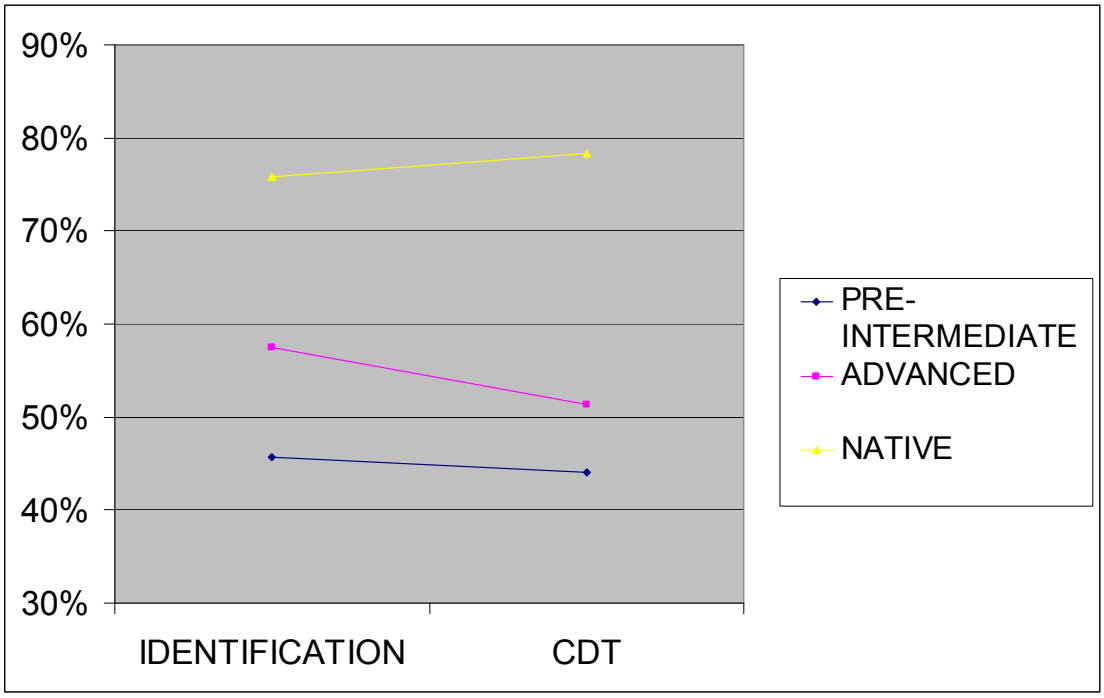

Figure 2. Performance on the perception tests by pre-intermediate learners, advanced learners and native speakers of English

As discussed in Kluge et al. (2007) the native speakers' results indicate that perception of final nasals may be somewhat complex even for native speakers. Even so, native speakers outperformed advanced learners who in turn outperformed pre-intermediate learners.

5. General discussion and conclusions

The question addressed in the present study was whether the adaptations of the CDT from phonemes to words – increasing the length of the trials – would compromise the results. A tentative answer from these results could be that WMC seems not to interfere in the scores of the CDT. Thus, the hypothesis was not supported and it indicates that results from Kluge et al.' adaptation of the CDT are not affected by WMC. Moreover, since a Pearson correlation yielded significant results for the CDT and the identification tests considering all groups, it seems that both perception tests are actually testing the same construct.

The hypothesis that WM capacity would affect results from the CDT was not supported. However, since only one version of the CDT and testing only one contrast (final /m-n/) was used, the results are tentative. Participants commented that while taking the CDT sometimes the previous trial (set of three words) interfered in what they remembered for the trial in question. The same pattern was found in the LST results. Sometimes, participants thought that a word from a previous set of sentences belonged to the current set. This

observation combined with what has been discussed on the load that the CDT poses to memory reinforced the hypothesis. Even so, non-significant, negative correlations were found between WMC and CDT's scores. Thus, it could be said that there may be an interference from WMC on CDT performance but it is rare and weak.

6. Limitations and suggestions for further research

Several drawbacks emerged from the present study and some of them are mentioned here. First, the sample was too small – the study could be replicated with a larger sample size. Second, the contrast between word-final nasals is difficult even for native speakers – perception tests assessing other non-native contrasts and adaptations of the CDT using phrases (as in Bettoni-Techio et al., 2007) and sentences (as in Silveiro, 2004) could be used. Finally, no definite answer for whether the nature of the relationship, if there is some, between WMC and perception assessed through adaptations of the CDT was provided.

Acknowledgments

The first author would like to thank the National Council for Scientific and Technological Development (CNPq) for partially supporting the present study. The second and third authors would like to thank the Committee for Postgraduate Courses in Higher Education (CAPES) for partially supporting the present study.

References

Bachman, L. F. &, Palmer, A. S. (1996). *Language testing in practice.* Oxford University Press.

Baddeley, A.D., & Logie, R.H. (1999). Working memory: The multiple-component model. In A. Miyake, & P. Shah (Eds.), *Models of working memory* (pp. 28–61). New York: Cambridge University Press.

Bettoni-Techio, M., & Koerich, R. D. (2004). A Percepção de alveolar plosivas finais na interfonologia de brasileiros aprendizes de Inglês. Unpublished paper. Universidade Federal de Santa Catarina, Florianópolis, Brazil.

Bettoni-Techio, M., Rauber, A. S., & Koerich, R. D. (2007). Perception and production of word-final alveolar stops by Brazilian Portuguese learners of English. In *Proceedings of Interspeech 2007* (pp. 2293-2296). Antwerp, Belgium.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin and Review*, *12*, 769-786.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behaviour*, *19*, 450-466.

de Fockert, J. W., Rees, G., Frith, C. D., & Lavie, N. (2001). The role of working memory in visual selective attention. *Science*, *291*, 1803-1806.

Flege, J. E. (1988). Factors affecting degree of perceived foreign accent in English sentences. *Journal of the Acoustical Society of America, 84*(1), 70-79.

Flege, J. E. (1995). Second Language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech Perception and linguistic experience: Issues in cross-language research* (pp. 233-272). Timonium, MD: York Press.

Flege, J. E. (1997). The role of phonetic category formation in second-language speech learning. In J. Leather, & A. James (Eds.), *New Sounds 97: Proceedings of the Third International Symposium on the Acquisition of Second Language Speech* (pp. 79-89). Klagenfurt, Austria.

Flege, J. E., Munro, M. J., & Fox, R. A. (1994). Auditory and categorical effects on cross-language vowel perception. *Journal of the Acoustical Society of America*, *95*(6), 3623-3641.

Fontanini, I., Weissheimer, J., Bergsleithner, J. M., Perucci, M., & D'Ely, R. (2005). Memória de trabalho e desempenho em tarefas de L2. *Revista Brasileira de Lingüística Aplicada*, *5* (2), 189-230.

Fortkamp, M. B. M. (2000). *Working memory capacity and L2 speech production: An exploratory study*. Unpublished doctoral dissertation, Universidade Federal de Santa Catarina, Florianópolis, Brazil.

Grossberg, S. (2003). Resonant neural dynamics of speech perception. *Journal of Phonetics, 31*, 423-445.

Guion, S.G., & Pederson, E. (2007). 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 honor of James E. Flege* (57-77). Amsterdam: John Benjamins.

Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition*, *14*, 25-38.

Keys, K. J. (2002). First language influence on the spoken English of Brazilian students of EFL. *ELT Journal*, *56*(1), 41-46.

Kluge, D. (2004). *Perception and production of final nasals by Brazilian Portuguese learners of English.* Unpublished Master's thesis. Universidade Federal de Santa Catarina, Florianópolis, Brazil.

Kluge, D. C., Rauber, A. S., Reis, M. S., & Bion, R.A.H. (2007). The relationship between perception and production of English nasal codas by Brazilian learners of English. In *Proceedings of Interspeech 2007* (pp. 2297-2300). Antwerp, Belgium.

Koerich, R. D. (2002). *Perception and production of word-final vowel epenthesis by Brazilian EFL students*. Unpublished doctoral dissertation, Universidade Federal de Santa Catarina, Florianópolis, Brazil.

Leather, J., & James, A. (1991). The acquisition of second language speech. *Studies in Second Language Acquisition*, *13*, 305-314.

Liberman, A., Harris, K., Hoffman, H., & Griffith, B. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, *54*, 358-368.

Miyake, A., & Friedman, N. P. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In A. F. Healy, & L. E. Bourne Jr (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 339-364). Mahawah, NJ: Lawrence Erlbaum.

Miyake, A., & Shah, P. (1999). Toward a unified theory of working memory: Emerging general consensus, unresolved theoretical issues, and future research directions. In A. Miyake, & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 442-481). Cambridge: Cambridge University Press.

Pennington, M. C. (1994). Recent Research in L2 phonology: Implications for Practice. In J. Morley (Ed.), *Pronunciation pedagogy and theory: New ways, new directions* (pp. 92-107). Illinois: Pentagraph Print.

Pisoni, D. B., & Tash, J. (1974). Reaction times to comparisons within and across phonetic categories. *Perception & Psychophysics, 15*, 285-290.

Rauber, A. S. (2004). *Working memory and the perception of English vowels by Brazilian Portuguese EFL speakers*. Unpublished paper, Universidade Federal de Santa Catarina, Florianópolis, Brazil.

Silveira, R. (2004). *The influence of pronunciation instruction on the perception and production of English word-final consonants*. Unpublished doctoral dissertation, Universidade Federal de Santa Catarina, Florianópolis, Brazil.

Silveiro, A. L. (2004). *The perception and production of English compound stress patterns by Brazilian learners of English*. Unpublished Master's thesis, Universidade Federal de Santa Catarina, Florianópolis, Brazil.

Tomitch, L. M. B. (1995). *Reading: Text organisation perception and working memory capacity*. Unpublished doctoral dissertation, Universidade Federal de Santa Catarina, Florianópolis, Brazil.

Torres, A. C. (2003). *Working memory capacity and reader's performance on main idea construction in L1 and L2*. Unpublished doctoral dissertation, Universidade Federal de Santa Catarina, Florianópolis, Brazil.