RESEARCH ON SPOKEN LANGUAGE PROCESSING
Progress Report No. 24 (2000)
Indiana University

The Influence of Short-term and Long-term Memory on the Identification and Discrimination of Non-native Speech Sounds [1]

James D. Harnsberger

Speech Research Laboratory
Department of Psychology
Indiana University
Bloomington, Indiana 47405

[1] This work was supported by NIH-NIDCD Training Grant DC00012 and NIH-NIDCD Research Grant DC00111 to Indiana University.

The Influence of Short-term and Long-term Memory on the Identification and Discrimination of Non-native Speech Sounds

Abstract. This study examined two possible sources of individual differences in cross-language speech perception, the capacity to phonologically encode speech and short-term memory span. Phonological coding was defined as the ability to encode non-native contrasts as distinct phonemes based on representations in long-term memory. Short-term memory was defined as a fixed capacity regulating the extent of encoded phonetic detail. To compare these two predictors of cross-language speech perception performance, thirty native speakers of American English were administered five tests: categorial AXB discrimination and identification (using non-native nasal consonant contrasts), digit span, nonword span (using pronounceable nonwords with nasal consonants, produced by a native speaker of English), and paired-associate word learning with word-word and word-nonword conditions. The AXB discrimination results were correlated with measures of short-term memory (digit span, word-nonword learning), phonological coding (identification), and a memory span measure mediated by phonological coding (nonword span). The results showed that almost all measures were significantly correlated with one another ($+0.62 > r > +0.41$), with the exception of word-word learning. The strongest predictor for the AXB discrimination test results was nonword span ($r = +0.62$, $p < 0.01$). When the identification test results were partialed out, only nonword span significantly correlated with discrimination ($r = +0.54$, $p < 0.01$). The results show an association between the discrimination of these non-native contrasts and a short-term memory capacity that interacts and relies heavily on prior linguistic experience in long-term memory.

Introduction

In numerous studies, a listener's prior linguistic experience has been shown to exert a profound effect on the identification, discrimination, and acquisition of non-native words (Flege, 1987; Miyawaki, Strange, Verbrugge, Liberman, Jenkins, & Fujimura, 1975). Specifically, listeners frequently identify and encode non-native sounds in terms of the most phonetically similar sounds in their native language. This encoding process can result in the loss of phonetic detail critical to differentiating non-native sounds. Such losses typically correspond to mismatches between the phonemic contrasts of the second language (L2) being acquired and those of the listeners' native language. For example, Japanese learners have particular difficulty in acquiring the /l/-/ɹ/ distinction in English, one which corresponds to a single Japanese liquid phoneme, usually realized as [ɽ] (Strange, 1995).

Such perceptual difficulties are not always explained, however, by a contrastive analysis of the phoneme inventories of the native and second languages in question. Listeners from the same linguistic background often vary substantially in the native sound(s) they use to encode a given non-native sound. For example, Schmidt (1996) found that Korean listeners used up to eight different consonants in labeling English /θ/ in an open-set identification test. In this case, the modal consonant used to label /θ/ only received 44% of the Korean listeners' responses. In a study of the identification of non-native nasal contrasts, Harnsberger (2000) observed highly variable identification patterns by seven groups of listeners for at least a subset of the non-native contrasts tested, particularly in the cases of native speakers of Malayalam, Punjabi, and Tamil. Such variable identification performance is not predictable by any type of contrastive analysis based on abstract units such as phonemes or allophones (Harnsberger, 2000).

Variability in the identification and discrimination of non-native sounds can often reflect large individual differences in perceptual performance. Such individual differences are rarely the central topic of cross-language speech perception studies, though they have been noted in passing (Bohn & Flege, 1990; Liberman, Harris, Kinney, & Lane, 1961; MacKain, Best, & Strange, 1981). In the literature on L2 acquisition, individual differences in learning are typically attributed to such factors as language aptitude, motivation, age of onset of learning, and the extent and type of experience in the second language, to name a few (Carroll, 1962; Carroll, 1981; Flege, Yeni-Komshian, & Liu, 1999; Miyake & Friedman, 1998; Skehan, 1989).

Of these factors, language aptitude is probably the most relevant to the problem of individual differences in cross-language speech perception, given that in such studies, listeners have no experience with the non-native words in question and are not attempting to acquire the language from which the non-native words were drawn. Language aptitude has been defined in past research on individual differences in L2 learning in terms of three component capacities: language analytic, phonetic coding, and short-term memory (Carroll, 1962; Carroll, 1981; Skehan, 1989). Language analytic capacity relates to the acquisition of L2 syntax. Phonetic coding refers to the ability to store non-native speech sounds in a manner that allows for easy storage and retrieval by the listener from long-term memory.

As a predictor of success in discriminating and identifying non-native sounds, phonetic (or, more accurately, phonological) coding can be conceived of as the capacity to detect and store in long-term memory the fine acoustic or articulatory details that differentiate the sounds and sound patterns of non-native words from similar sounds and patterns used in native words. Short-term memory capacity refers to a memory component of fixed capacity whose purpose is to temporarily store information for encoding in long-term memory. Traditionally, short-term memory has been modeled as a general mechanism that is independent of representations in long-term memory. For example, Baddeley's (1986) model of working memory includes the phonological loop, which combines a phonological short-term store of fixed capacity with a subvocal rehearsal process to maintain the phonological representation. The short-term store in this model has been modeled as independent of any input from representations in long-term memory, which would include the listener's prior linguistic experience (though see Baddeley, Gathercole, & Papagno, 1998; Carpenter, Miyake, & Just, 1994).

The evidence supporting phonological coding as the basis of individual differences in cross-language speech perception can be found primarily in the repeated demonstration of a relationship between the identification of non-native sounds and their discrimination (see Strange, 1995 for a review of the cross-language speech perception literature). Identification of non-native sounds with native phonemes has also been shown to extend to L2 acquisition, both in perception (Miyawaki et al., 1975) and production (Flege, 1987). While extensive experience can allow listeners to form new perceptual categories for non-native sounds (Lively, Logan, & Pisoni, 1993; Lively, Pisoni, Yamada, Tohkura, & Yamada, 1994; Logan, Lively, & Pisoni, 1991; MacKain et al., 1981), in some instances, even highly experienced learners may fail to correctly identify non-native sounds (Flege, 1991).

The evidence suggesting a role for short-term memory capacity in cross-language speech perception comes from studies of novel word learning by children, adults with cognitive deficits, and normal adults. Like cross-language speech perception, novel word learning involves the perception and encoding of novel phonological forms. Numerous studies have demonstrated significant correlations between short-term memory span and novel word learning in children (Gathercole, 1995; Gathercole & Baddeley, 1990; Gathercole, Frankish, Pickering, & Peaker, 1999; Gathercole, Service, Hitch, Adams, and Martin, 1999; Gathercole, Willis, Emslie, & Baddeley, 1991). Selective impairment in short-term memory capacity has also been shown to affect nonword learning in paired-associate word-learning tasks

(Baddeley, Papagno, & Vallar, 1988; Papagno & Vallar, 1995a; Trojano & Grossi, 1995). This relationship has been observed in normal adults as well (Papagno, Valentine, & Baddeley, 1991; Papagno & Vallar, 1992; Service, 1998).

Recently, research on short-term memory has also been extended to L2 learning by children and adults. Papagno et al. (1991) examined non-native (Russian) word learning by two groups, Italian and English native speakers, under conditions of articulatory suppression. Articulatory suppression involves the repetition of a speech sound or word while engaging in a concurrent task that relies on short-term memory. The effect of articulatory suppression is to interfere with the maintenance of phonological information in short-term memory. Papagno et al. predicted that articulatory suppression would disrupt novel word learning but not paired-associate word-word learning, assuming novel word learning relies heavily on short-term memory. The participants were tested in a paired-associate word-learning task with word-word and word-nonword pairs. Both the English and Italian speakers displayed the predicted pattern of poor performance on word-nonword pairs.

Service (1992) conducted a longitudinal study of English acquisition by nine year-old children who were native speakers of Finnish. Service found significant correlations between three nonword repetition tasks and grades in English and in other classes. MacKay, Meador, and Flege (2000) observed modest correlations between a nonword repetition measure and the error rates in identifying English word-initial and word-final consonants by native speakers of Italian. Miyake and Friedman (1998) studied the acquisition of English syntax by 59 native speakers of Japanese. Subjects were administered five tests, two on syntactic comprehension, two listening span measures using Japanese and English stimulus materials, and a digit span task. Listening spans were considered to be measures of "working memory" rather than short-term memory tests, such as digit span, because listening spans incorporate a computational or operational component. Miyake and Friedman found that the L2 working memory span measure (with English materials) correlated significantly with the syntactic comprehension measures and the L1 working memory span measure (with Japanese materials). Interestingly, the digit span measure did not correlate significantly with the syntactic comprehension measures. Thus, the strongest correlations observed in this study involved measures that used a similar type of stimuli (i.e., sentences in the working memory and syntactic comprehension measures), indicating an effect of long-term memory on representations in short-term memory.

The effect of short-term memory on L2 acquisition has also been studied indirectly by examining individual differences in multilingual, bilingual and monolingual populations. Papagno and Vallar (1995b) compared ten "polyglot" and ten "nonpolyglot" speakers in a number of tasks, including general intelligence, vocabulary (i.e., a subtest of Wechsler Adult Intelligence Scale), auditory digit span, nonword repetition, visual-spatial span, visual spatial learning, and paired-associate word-nonword learning. Polyglots were defined as speakers of at least three languages, while nonpolyglots were defined as speakers or two or fewer languages. The two groups' performance differed significantly only on the auditory digit span, nonword repetition, and the word-nonword condition of the paired-associate word-learning task. Papagno and Vallar argued that a large phonological short-term memory capacity aids in the acquisition of novel words. Thus, individuals with a greater capacity in phonological short-term memory are predicted to be better at acquiring multiple languages.

While a number of studies have shown a relationship between short-term memory and novel word learning in several populations of interest, several questions remain concerning the methodologies of these studies and the relationship between short-term and long-term memory, particularly as they apply to the problem of individual differences in cross-language speech perception. First, many of these studies have relied on nonword repetition as the primary measure of memory span. The use of this methodology may be particularly problematic with certain populations, such as children, adults with particular cognitive deficits, or adults attempting to produce non-native (unfamiliar) sounds, because it confounds short-term memory capacity with speech production skills. This issue was addressed recently by Gathercole, Service, Hitch, Adams, and Martin (1999). In their study, eighteen four-year-old children were administered digit span, nonword repetition, vocabulary, and nonword matching span tasks. Most importantly, the nonword matching span tasks did not require verbal responses. Nonword matching span was found to correlate significantly with vocabulary, demonstrating that with at least this population, nonword repetition is a valid measure of memory span.

Second, the few studies of short-term memory and second language acquisition by normal adults and children have employed pronounceable nonwords in tasks such as nonword repetition or paired-associate word learning (Papagno & Vallar, 1995b; Thorn & Gathercole, 1999). Obviously, it would be difficult to use stimulus materials such as non-native words that are difficult to accurately produce in tests requiring a verbal response. However, the learning of non-native words that are difficult to accurately identify and/or produce is a common problem faced by learners of many languages. Acquiring receptive and expressive mastery of spoken words differentiated by phonemic contrasts that are difficult to distinguish constitutes a significant learning problem for adult learners of second languages (Flege, 1992). Examining the role that short-term memory plays in the acquisition of non-native words requires procedures that do not involve verbal responses (such as a matching span task) or correlations of span measures with identification or discrimination tests using non-native words.

Finally, memory span tests often confound possible individual differences in memory capacity with individual differences in the perception, encoding, and representation of linguistic knowledge in long-term memory. Recent work has shown that phonological short-term memory measures vary greatly depending on a listener's prior experience with the stimulus materials used. For example, "wordlikeness," or the similarity of a nonword to a real word, has been shown to effect short-term memory spans of adult speakers of English (Papagno et al., 1991) and English-speaking children (Gathercole, 1995). Short-term memory spans of children have also been shown to be influenced by the phonotactic properties of the stimulus materials (Gathercole, Frankish, Pickering, & Peaker, 1999). More importantly, whether the stimulus materials consist of native or non-native sounds can also affect subjects' performance on short-term memory tasks. Thorn and Gathercole (1999) found that English-French bilingual children and English-speaking children learning French as a second language had significantly longer French nonword repetition spans than monolingual English children. All of these studies strongly suggest that phonological coding in short-term memory, rather than being "knowledge-free" or independent of knowledge and representations in long-term memory, relies heavily on prior knowledge, specifically, on linguistic experience.

A close connection between short-term and long-term memory poses problems for theories of language acquisition that claim that fixed memory capacities in individuals affect the learning of new words or new grammatical rules. Individuals may also differ in the nature of their long-term experience with speech that resembles the stimulus materials used in memory span tasks. Such differences in prior linguistic experience may manifest themselves as differences in vocabulary size. However, individuals may also differ in the extent of their experience with different phonological forms of different words in the lexicon. Listeners with more extensive listening experience with multiple languages, or in multiple dialects of their native language, may possess word representations that incorporate a greater range of acoustic-phonetic detail and variability. Numerous earlier studies have clearly shown that listeners encode precise details of episodes of speech (see Goldinger, 1998 for a review). Such "robust" episodic word representations may aid listeners in the performance of tasks such as memory span tests. Thus, memory span, rather than being an important source of individual differences in language acquisition, may be the "byproduct" of individual differences in prior linguistic experience. It is also possible that both fixed short-term memory capacity and prior linguistic experience may interact to determine the capacity of listeners to store phonetically detailed representations of non-native words for encoding in long-term memory (see Baddeley et al., 1998 for an example of a word-learning model in which short-term and long-term memory interact).

The purpose of this study was to examine the effects of short-term memory capacity and prior linguistic experience on individual differences in cross-language speech perception by native speakers of American English. Short-term memory capacity was measured independently of long-term memory by using an immediate serial recall task with highly familiar stimulus materials, namely, English digits. In addition, a paired-associate word-learning task, using word-word and word-nonword pairs, was administered. The word-word pairs served to separate the "pure" short-term memory span task with digits from any effects of individual differences in familiar word learning. The word-nonword learning task was used as an analog to acquiring unfamiliar words in a second language and as a language learning measure strongly related to short-term memory span (Papagno & Vallar, 1995b).

Prior linguistic experience was measured indirectly by measuring the phonological coding capacity of listeners. Phonological coding, as described by Skehan (1989), corresponds to a listener's reliance on long-term memory representations based on prior linguistic experience. This ability was measured using an identification test with non-native words that were thought to be difficult to discriminate for English listeners. To compare the relative success of the short-term memory and phonological coding measures in predicting the cross-language speech perception abilities of English listeners, a discrimination test using the same non-native words as the identification test was administered. Discrimination tests are commonly employed in cross-language speech perception research and do not require an overt verbal response. The possible interaction of pure short-term memory capacity with prior linguistic experience was measured with an immediate serial recall memory span task using native nonword analogs to the non-native words under study. If prior linguistic experience influences short-term memory capacity, then native nonword spans should show stronger correlations with the discrimination test results than either the identification, digit span, or word-nonword-learning test results.

If an effect of short-term memory were observed on the discrimination test results, then an important source of individual differences in cross-language speech perception would be identified. Currently, models of cross-language speech perception, such as the Perceptual Assimilation Model (Best, 1995) or the Native Language Magnet model (Kuhl & Iverson, 1995), do not formally incorporate a role for individual differences in short-term memory in the discrimination of non-native contrasts. In both models, discrimination is a function of the manner in which the sounds constituting a non-native contrast (e.g., dental and retroflex stops for English listeners) are identified with, or phonologically coded as, one or more native sounds. In the Native Language Magnet model, the identification-discrimination relationship is described in terms of the association of stimuli to particular locations in perceptual space, within a particular perceptual category. Two non-native stimuli that constitute a non-native contrast are discriminable based in their locations in perceptual space: Are they close to a prototype, or the category periphery? In the Perceptual Assimilation Model, the identification-discrimination relationship is also a function of the proximity of the stimuli to one or more perceptual categories: Are the stimuli equally good exemplars of a single category, do they differ in category goodness, or are they even associated with any native perceptual categories?

In both of these models, the preservation of phonetic detail in the encoding of a non-native contrast in long-term memory is the automatic consequence of how a contrast was phonologically coded in terms of one or more existing representations in long-term memory (i.e., perceptual categories). However, both models would require revision if the results of this study demonstrate a role for short-term memory in either the discrimination or the identification of non-native sounds. The results of this study may demonstrate that an individual's short-term memory capacity constrains the encoding of phonetic detail in non-native sounds, independent of the individual's phonological coding, or identification, of non-native sounds. Alternatively, short-term memory capacity itself may determine individual patterns in the identification and discrimination of non-native sounds. Finally, short-term memory capacity may prove to be unrelated to cross-language speech perception.

Main Experiment

Methods

Participants

Thirty normal-hearing Indiana University undergraduates participated in this experiment, 11 males and 19 females. The participants ranged in age between 18 and 27 (M = 22). No subject reported any history of a speech or hearing problem. For participating in two 1-hour sessions, the subjects received $15 as compensation.

Stimulus Materials

AXB Discrimination and Identification Tests. The stimulus materials used for these tests were a subset of those used by Harnsberger (2000). Briefly, one male and one female talker of Malayalam, a Dravidian language of southern India, were recorded reading from a list of real and nonsense words from their native language that included all six nasal consonants of Malayalam. The nasals of interest appeared in all syllable positions allowable by the individual languages, in an /a/, /i/, or /u/ vocalic context. All of the stimuli recorded were evaluated by native speakers of the respective languages in an identification test in order to exclude stimuli that might be poor exemplars of Malayalam nasals. Of the stimuli that were recorded and evaluated, a subset was used in these tests: four exemplars from an isolation context, two from each talker, of dental, alveolar, and retroflex medial geminate nasals in an [i] context. These three types of nasal consonants were selected because they could be combined to form non-native contrasts that have been shown to be of intermediate difficulty for English listeners to discriminate (Harnsberger, 1998).

In addition to the non-native words, a similar set of nonsense words were recorded from a male speaker of American English, who produced two exemplars each of three nasal consonants, [m], [n], and [ŋ], appearing in final position following [i]. The native nonsense words with bilabial and alveolar nasals were paired together to form AXB discrimination test trials. All of the native nonsense words were included in the identification test. All of the native and the non-native nonsense words were leveled in amplitude at 70 dB for presentation in the experiment.

Digit Span. The stimulus materials for the digit span test included one exemplar each of ten words, "one," "two," "three," "four," "five," "six," "seven," "eight," "nine," and "zero," produced by a male speaker of English. The stimuli were leveled in amplitude at 70 dB for presentation in the experiment.

Nonword Span. The stimulus materials for the nonword span test were a subset of the native nonsense words used in the identification and discrimination tests: one token each of the bilabial, alveolar, and velar native nonsense words. The stimuli were leveled in amplitude at 70 dB for presentation in the experiment.

Paired Associate Word Learning. The stimulus materials for the paired-associate word-learning test consisted of eight pairs of real words and eight pairs of pronounceable nonwords, all produced by a male speaker of American English. The sixteen pairs are listed in the Appendix. The real word pairs consisted of four two-syllable and four three-syllable pairs, consisting of high frequency, concrete nouns. Words appearing in pairs were matched together in terms of their stress pattern. The nonsense word pairs also consisted of four two-syllable and four three-syllable pairs. The nonsense words were created from syllables involving highly probable phoneme sequences (Vitevitch, Luce, Charles-Luce, & Kemmerer, 1997). The stimuli were leveled in amplitude at 70 dB for presentation in the experiment.

Procedure

Five tests were administered to the subjects over two sessions. In the first session, subjects were administered an AXB discrimination test and an identification test. The order in which these tests were administered was counterbalanced. In the second session, the subjects were administered the span tests (digit span first, nonword span second), and the paired-associate word-learning test (word-word pairs first, word-nonword pairs second). The span tests were counterbalanced with the paired-associate word-learning tasks.

Identification Test. The identification test was a forced-choice identification test consisting of 90 trials, specifically, five repetitions of eighteen nonwords. Each word appeared in isolation, and the order of presentation of trials was randomized. Listeners were instructed to label the nasal consonant of each word using one of three response choices (keywords) corresponding to the three nasal phonemes of English, “sum” (/m/), “sun” (/n/), or “sung” (/ŋ/). Three familiarization trials, consisting of three native nonsense words ([im], [in], and [iŋ]), were presented to confirm that listeners understood which nasal consonants were represented by the keywords.

In scoring the identification test results, subject responses to pairs of stimulus types that corresponded to discrimination test trials were compared. For instance, the identification test results of the alveolar and retroflex nasals produced by Malayalam talker YM were scored together to yield a predicted discrimination test score for the alveolar-retroflex contrast produced by YM that appeared in the discrimination test. The scoring method involved calculating the extent to which two stimulus types (e.g., bilabial and alveolar nasals) differ in their identification with native perceptual categories, a metric termed hereafter the *categorization difference score* (C-score). Specifically, the C-score is based on the sum of the differences between the two stimulus types in their proportion of responses for each native perceptual category. The C-score is represented in equation (1):

$$C = \frac{\sum_{i=1}^{n} |A_i - B_i|}{2t} \quad (1)$$

where, C is the categorization difference score, n is the number of response categories available to the listener, A_i is the number of responses of stimulus type A to category i, B_i is the number of responses of stimulus type B to category i, and t is the number of trials in which each stimulus type was presented (assuming each stimulus type was presented the same number of times over the course of the identification test). $2t$ is a constant that simply converts the range of scores to a familiar 0.0 – 1.0 scale.

This metric was used by Harnsberger (1999) successfully to account for differences in the discriminability of a large set of non-native contrasts by seven different listener groups in a cross-language speech perception study. The calculation of a C-score can be illustrated using the hypothetical data set shown in Table 1. In this example, the proportion of responses to each stimulus type ([n̪], [n], [ɳ]) is listed below each response choice (/m/, /n/, or /ŋ/). In this example, each stimulus type was presented ten times over the course of the identification test. To calculate the C-score for the [n̪]-[n] contrast, the absolute value of the difference between the proportion of /m/ responses to [n̪] and [n], 8 – 0, must be summed with the absolute value of the difference between the proportion of /n/ responses to [n̪] and [n], 0 – 9, and the absolute value of the difference between the proportion of /ŋ/ responses to [n̪] and [n], 2 – 1. The resulting value of 18 (|8 – 0| + |0 – 9| + |2 – 1|) is divided by two times the number of trials in which each stimulus type was presented (2 * 10 = 20) to yield a relatively high C-score of 0.9. In this example, the C-score for the [n̪]-[ɳ] contrast is a relatively low 0.3 ({|8 – 5| + |0 – 3| + |2 – 2|}/20), while the C-score for the [n]-[ɳ] contrast is an intermediate 0.6 ({|0 – 5| + |9 – 3| + |1 – 2|}/20).

	Response choices		
Stimulus Type	/m/	/n/	/ŋ/
[n̪]	8	0	2
[n]	0	9	1
[ɳ]	5	3	2

Table 1. A hypothetical dataset for the identification test. The number in each cell is the number of trials in which a particular response choice (e.g., the English bilabial nasal represented by the keyword "sum") was selected for a particular stimulus type (e.g., a dental nasal).

C-scores were used to predict the discrimination test scores. Contrasts involving stimulus types that were frequently labeled in a different manner (corresponding to high C-scores) were predicted to be relatively more discriminable than contrasts involving two stimulus types that frequently received the same label (corresponding to low C-scores). In this study, the C-scores tested the phonological coding hypothesis in a correlation analysis with the other test scores.

Discrimination Test. The discrimination test was a categorial AXB test consisting of 112 trials, 16 trials each of the seven different types of contrasts, where "contrast" refers to a particular place distinction produced by a particular talker. Six of these contrasts were non-native, namely, dental-alveolar, dental-retroflex, and alveolar retroflex produced by the two Malayalam talkers. One contrast was produced by the American English talker, namely, bilabial-alveolar. This native contrast was included as a control, to test whether individual discrimination scores varied due to the difficulty of the stimuli as opposed to any difficulties with the test format.

In order to ensure that the results were not dependent on the intelligibility of a single exemplar, two exemplars of each member of the seven contrasts were used. The contrasts appeared in four possible orders, AAB, ABB, BAA, and BBA. A and B were always from the same talker, and all stimuli that were paired together were selected to minimize acoustic differences that were not relevant to the identity of the stimulus, such as the overall duration or the fundamental frequency pattern of a stimulus. The interstimulus interval for the discrimination test was 1 s. The order of presentation of AXB trials was also randomized.

Subjects were told to indicate whether the nasal consonant in the first or last word was the same as the nasal consonant in the middle word by circling a number on the answer sheet. The term "nasal consonant" was defined through the use of simple examples in which nasals appeared in different syllable positions and vocalic contexts. A, X, or B were not physically identical, so listeners made categorial matches. One familiarization trial was presented before the AXB test. The discrimination test results were analyzed in terms of mean percent correct responses for individual contrasts as well as a mean score over all contrasts.

Digit Span. In the digit span test, subjects were presented auditorily with a sequence of single digits (0-9) pronounced by a native speaker of American English. Subjects were asked to write down in order the digit sequence presented. The length of the digit sequences began at four, and ended at ten, with two trials at each sequence length in order of increasing length. The results of the digit span task were scored in two ways: first, in terms of the longest sequence length in which a subject correctly recalled all digits in order in both trials; second, in terms of an *absolute span score* which sums each correct trial weighted by its sequence length, as in equation (2):

$$ABS = \sum_{i=1}^{n} x_i * y \qquad (2)$$

where, *ABS* is the absolute span score, *i* is a trial, *n* is the total number trials, *x* is the length of the digit sequence of trial *i*, and *y* is the accuracy of the subject's response (correct = 1, incorrect = 0). Both the longest digit span score and the absolute digit span score were used in a correlation analysis with the results of the other tests.

Nonword Span. In the nonword span test, subjects were presented auditorily with a sequence of nonwords consisting of VC syllables ([im], [in], [iŋ]) pronounced by a native speaker of American English. Subjects were asked to write down in order the nonword sequence presented using the symbols "m," "n," and "ng" for [im], [in], and [iŋ], respectively. The length of the nonword sequences began at three, and ended with seven, with two trials at each sequence length in order of increasing length. The results of the nonword span task were scored in terms of the longest sequence length in which a subject correctly recalled all nonwords in order in both trials and in terms of an absolute span score, described earlier. As with the digit span test, whichever of the two scoring methods showed the strongest correlation with the discrimination test scores was used in subsequent analyses.

Paired Associate Word Learning. This paired-associate word-learning test consisted of two conditions, word-word and word-nonword. In the word-word condition, subjects were presented with the full set of word-word pairs. Each word in the pair was separated by a 1 s interval, while each pair was separated by a 2 s interval. After hearing all of the word-word pairs, subjects were presented with the first word from each pair and responded verbally with its corresponding second word, if it could be recalled. This second part of the procedure was self-paced. Subjects received a "correct" score for each trial if their response matched the second word exactly. This entire sequence (presenting the word-word pairs followed by the first words from each pair) was repeated until the subject recalled all the second words correctly in two consecutive repetitions of the entire sequence, or until the tenth repetition was complete. A different random trial order was used for each repetition. The procedure for the word-nonword condition was identical to the word-word condition, except for the stimuli presented to the subjects. The Appendix lists the particular word-word and word-nonword pairs used in this test. The word-word and the word-nonword conditions were scored in terms of the number correct of correct responses at a particular repetition in the condition.

Predictions

Three hypotheses were entertained in this study. First, the short-term memory hypothesis states that listeners' fixed short-term memory capacity determines the extent of phonetic detail that is encoded in representations in long-term memory. Thus, greater short-term memory capacities should make nonword learning easier, resulting in a significant correlation between digit span and word-nonword learning. In addition, short-term memory should also determine the manner in which non-native words are phonologically encoded. Thus, digit span should correlate significantly with the identification and discrimination test scores. Any significant correlations between nonword span and the identification and discrimination tests were predicted to disappear if digit span was partialed out.

Second, the phonological coding hypothesis states that the extent of phonetic detail that is encoded in long-term memory representations of non-native or novel words is determined by the manner in which the words are phonologically encoded, that is, by the native sounds used to encode the nonwords. According to this hypothesis, an individual's manner of encoding is based his/her prior linguistic experience rather than his/her short-term memory capacity. If this is the case, then the identification test results should correlate with the discrimination test results. In addition, no significant correlation was predicted between either the discrimination or identification test results and the digit or nonword spans.

Finally, the short-term/long-term memory (STM-LTM) interaction hypothesis states that the extent of encoded phonetic detail is determined by a short-term memory capacity that receives input and is influenced by linguistic representations in long-term memory. Thus, when listeners encode speech in short-term memory, they rely on representations in long-term memory, though the capacity of their short-term memory should still constrain their ability to accurately encode phonetic detail in non-native words. According to this hypothesis, the discrimination and identification results should correlate with the various span measures, but nonword span should correlate more strongly than the pure measure of short-term memory, digit span. That is, individuals with long-term memory representations that incorporate phonetic details that aid them in encoding these non-native words should also perform better on span tasks that employ stimulus materials similar to those of the discrimination and identification tests (i.e., nonword span), relative to span tasks that do not utilize those long-term memory representations (i.e., digit span).

Results

Scores for Individual Tests

The results of the discrimination test (AXB) as well as the identification test (ID) are shown in Table 2. The results are listed by contrast, defined in terms of place of articulation as well as talker. The AXB results are the mean proportion of correct responses, while the ID scores are the mean C-scores (the standard deviations of all means are given in parentheses). The native contrast elicited a relatively high mean percent correct score of 89%, though subjects did not perform at ceiling as might have been expected. Across all non-native contrasts, subjects averaged 0.65 proportion correct (SD = 0.18) in the discrimination test and a 0.38 C-score (SD = 0.28) in the identification test. These discrimination test scores are somewhat lower than those elicited from native speakers of American English with the same stimuli in a previous study by Harnsberger (1998): the dental-alveolar, dental retroflex, and alveolar-retroflex contrasts of Malayalam talker YM elicited mean proportion correct discrimination scores of 0.9, 0.8, and 0.83, while the same three contrasts from talker YS elicited scores of 0.59, 0.68, and 0.64.

Place Contrast	Talker	AXB	ID
bilabial-alveolar	JH	0.89 (0.12)	0.87 (0.23)
dental-alveolar	YM	0.8 (0.17)	0.31 (0.31)
	YS	0.57 (0.15)	0.2 (0.15)
dental-retroflex	YM	0.71 (0.21)	0.45 (0.34)
	YS	0.52 (0.11)	0.34 (0.2)
alveolar-retroflex	YM	0.71 (0.13)	0.54 (0.29)
	YS	0.62 (0.13)	0.41 (0.19)

Table 2. The discrimination test results reported as mean proportion of correct responses and listed by contrast (place and talker). The standard deviations of the means are in parentheses.

The distribution of individual scores on both tests is shown in Figure 1, organized by talker (YM and YS) and by place contrast (dental-alveolar, dental-retroflex, and alveolar-retroflex). Across all contrasts, a sufficient spread in the discrimination and C-scores were observed, insuring that the cross-language test results could be used in correlation analyses. In particular, the results for the contrasts produced by YM elicited a great range of results, as indicated by high standard deviations and relatively shallow slopes in the bell-shaped curves of the distributions.

The digit span and nonword span scores also showed significant variability in the individual results. The mean digit span scores, measured in terms of longest span correct and absolute span, were 6.9 (SD = 1.6) and 51.8 (SD = 20.9), respectively. The mean nonword spans were 4.6 (SD = 1.8) and 21.8 (SD = 13.4), for the longest span and absolute span scores, respectively.

Figure 2 shows the number of correctly recalled second words for each repetition in the paired-associate word-learning test, for both the word-word condition (solid line) and the word-nonword condition (dashed line). The word-word pair learning proved to be a relatively easy task for the subjects of this experiment. By the fourth repetition, subjects were averaging close to ceiling-level performance (7.7 words). The word-nonword pairs proved to be more difficult to learn than the word-word pairs, as expected. However, the rate of learning in each condition was similar, though subjects on average did not asymptote at ceiling with the word-nonword pairs. Given the results of this test, the mean number of correctly recalled words from the first repetition of the word-word condition was taken to represent the performance of individual subjects. For the word-nonword condition, the mean number of correctly recalled words from the fifth repetition was taken as the score. These two repetitions were chosen because both elicited intermediate values in the 0 – 8 scale of the test, avoiding any ceiling or floor effects in comparing the paired-associate word-learning results to those of other tests.

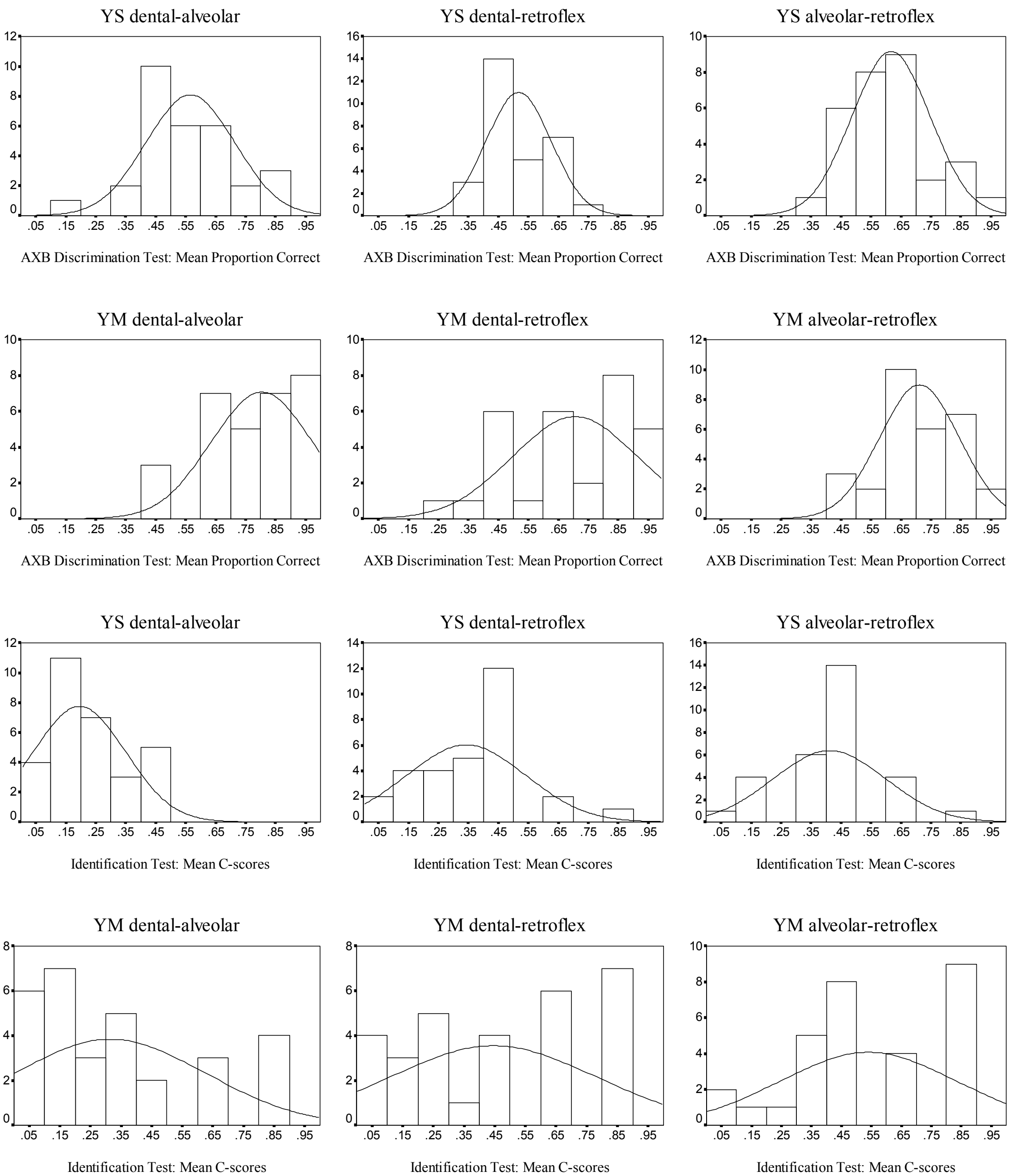

Figure 1. Histograms of the distributions of individual scores on the AXB discrimination and identification tests.

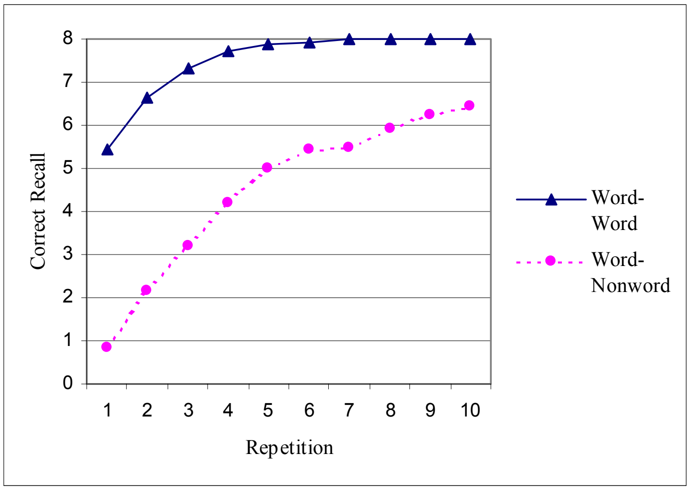

Figure 2. The number of correctly recalled words or nonwords for each repetition of the word-word or word-nonword pairs.

Full Correlation Matrix

Table 3 shows the full correlation matrix of the results, including the discrimination test scores (AXB), the categorization difference scores from the identification test (ID), the digit span test scored in absolute span (DigABS) as well as longest span (DigLong), the nonword span test scored in absolute span (NonABS) as well as longest span (NonLong), the mean number of correct words in the first repetition of the word-word condition of the paired-associate word-learning test (WW), and the mean number of correct words in the fifth repetition of the word-nonword condition (WN). In correlating the discrimination test scores with the scores for other tests, mean percent correct scores that were averaged over all contrasts presented in the test were used, with one exception: the coefficient for the AXB – ID correlation represents an analysis in which the discrimination and C-scores of individual contrasts were entered, since unique C-scores were available for the individual contrasts.

Test	AXB	ID	DigAbs	DigLong	NonABS	NonLong	WW	WN
AXB	X							
ID	0.41**	X						
DigAbs	0.34	0.53**	X					
DigLong	0.3	0.42*	0.93**	X				
NonABS	0.62**	0.44*	0.38*	0.39*	X			
NonLong	0.57**	0.34	0.34	0.39*	0.91**	X		
WW	0.25	0.16	0.33	0.2	0.22	0.26	X	
WN	0.53**	0.41*	0.51**	0.43*	0.46**	0.44*	0.29	X

Table 3. The full correlation matrix for all tests administered in the experiment, including discrimination (AXB), identification (ID), digit span (DigABS and DigLong), nonword span (NonABS, NonLong), and the word-word (WW) and word-nonword (WN) conditions of the paired-associate word-learning test. * $p < 0.05$, ** $p < 0.01$

According to the short-term memory hypothesis, the discriminability of non-native or novel words should be a function of an individual's fixed capacity to encode phonetically-detailed representations in short-term memory for storage in long-term memory. In contrast, the phonological coding hypothesis states that individuals can vary greatly in how they identify, or encode, non-native sounds in short-term memory, and those differences are a function of the unique properties of their perceptual or lexical categories in long-term memory (i.e., the region in multidimensional acoustic space that the category occupies). Finally, the STM-LTM hypothesis stated that the discriminability of novel and non-native words is a function of both an individual's phonological coding strategy and a fixed capacity to encode the phonetic details of novel words.

The results in Table 3 indicate that a traditional method of measuring pure short-term memory, digit span, failed to correlate significantly with the discrimination test results, regardless of the scoring method used ($r = +0.3$ for longest span, $r = +0.34$ for absolute span). Phonological coding, as represented by C-scores, did correlate significantly ($r = +0.41$, $p < 0.01$), although the strength of the correlation was not as great as that observed by Harnsberger (1999), who examined a much larger data set using the same scoring method. However, the strongest correlations with the discrimination test results were found with the memory span measures using nonword stimulus materials. The word-nonword learning scores correlated significantly with the discrimination test scores ($r = +0.53$, $p < 0.01$), while the nonword spans, measured using stimulus materials that were quite similar to those of the discrimination test, showed the strongest correlation ($r = +0.62$, $p < 0.01$, for the absolute span scoring method).

Taken alone, the correlations with the discrimination test support the STM-LTM interaction hypothesis, which predicted that the span measure employing the most similar stimuli to those of the identification and discrimination tests would show the strongest correlation. However, several other measures also correlated with the results from tests using the non-native stimuli. While pure short-term memory (as represented by the digit span task) did not correlate with discrimination, it did correlate significantly with the identification test results, the nonword spans (when matched in scoring method), and the word-nonword scores. In fact, the identification, discrimination, digit span, and nonword measures were all intercorrelated. Only the word-word scores failed to significantly correlate with any of the other measures in the analysis, suggesting that some form of memory span, as opposed to a general ability in verbal learning, accounts for the individual differences in the perception tests with non-native stimuli (see Papagno and Vallar, 1995b, for a similar finding on paired-associate word learning and short-term memory tasks).

Partial Correlations

To examine further the factors responsible for individual differences in cross-language speech perception, a number of partial correlations were computed to separate the effects of short-term memory and long-term memory on listeners' perceptual performance with non-native words. The results shown in Table 3 failed to support the phonological coding hypothesis. The remaining hypotheses were differentiated by partialing out the effect of short-term memory (as indexed by digit span). If short-term memory capacity alone is responsible for individual differences in the encoding of the phonetic details of non-native words, then in this partial correlation analysis, nonword spans should not significantly correlate with identification and discrimination test results. Only the digit and nonword spans scored using the absolute span score method were used in this analysis because in the original correlations, digit and nonword spans showed the strongest correlations with other measures when scored by that method. In addition, the word-word learning results were dropped from the analysis, since they did not correlate with any of the other measures obtained.

The resulting correlation matrix with digit span partialed out appears in Table 4. In this analysis, the identification and word-nonword scores no longer correlate with the discrimination scores. However, the strong correlation between discrimination and nonword spans still remained. When partialing out all other measures, nonword span was still significantly correlated with the discrimination scores ($r = +0.46$, $p < 0.05$). The importance of nonword span for discrimination was demonstrated again in a third correlation analysis in which nonword span was partialed out. In this analysis, no other factor correlated significantly with discrimination; digit span, however, was still correlated with identification ($r = +0.43$, $p < 0.05$) and word-nonword learning ($r = +0.41$, $p < 0.05$).

Test	AXB	ID	NonABS	WN
AXB	X			
ID	0.26	X		
NonABS	0.56**	0.3	X	
WN	0.44*	0.19	0.34	X

Table 4. The correlation matrix with digit span (DigABS) partialed out, including discrimination (AXB), identification (ID), nonword span (NonABS), and the word-nonword (WN) condition of the paired-associate word-learning test. * $p < 0.05$, ** $p < 0.01$

Discussion

Of the three proposed hypotheses, the results of this study support the STM-LTM interaction hypothesis: phonological short-term memory, as measured using native stimulus materials that were phonologically similar to the non-native words under study, correlated significantly with the results of cross-language discrimination and identification tests. This effect of phonological similarity would only be observed if the short-term memory of listeners is influenced by long-term memory representations of words involving nasal consonants. Thus, long-term memory plays an important role in the encoding of non-native contrasts in short-term memory.

However, short-term memory capacity is not simply a byproduct of variation in individual differences in long-term memory representations. Digit spans, a measure using highly familiar stimulus materials that is assumed to tap pure short-term memory, correlated significantly with many of the measures collected here, including word-nonword learning and, most importantly, identification. The correlation between digit span and identification, but not discrimination, indicates that identification and discrimination abilities are separable to some extent, and may rely on short-term and long-term memory differently. Nevertheless, nonword span showed the strongest correlations with the discrimination test scores, and correlated with nonword learning and identification in a similar manner as digit span. The strength of the nonword span correlations supports the STM-LTM interaction hypothesis, which has also received support in several other recent studies in which either the stimulus materials or listener groups are varied linguistically (Gathercole, Frankish, Pickering, & Peaker, 1999; Thorn & Gathercole, 1999).

The success of the nonword spans, compared with digit spans, as a predictor of cross-language discrimination implies that as the similarity between the stimulus materials of the span task and those of a correlated measure (such as discrimination) increases, the strength of the correlation increases. That is, the effect of long-term memory on span capacity should be greater for stimulus materials that are more similar to representations in long-term memory. This proposed relationship can be tested by examining a subset of the data collected in this study. While the stimulus materials of the non-word span test were phonologically similar to the non-native stimuli, they were most similar to the native control stimulus materials of the identification and discrimination tests. Specifically, they were a subset of the native nonsense words included in the identification test. The discrimination test used a different subset of these nonsense words: two tokens each of the nonsense words with bilabial and alveolar nasals. The same male talker produced all of the native nonsense words used in this study, and all of the nonsense words were of the form [iN], where N is a nasal consonant of English, either [m], [n], or [ŋ]. Given the similarity of the native stimulus materials of the identification and discrimination tests and those of the nonword span task, we can predict that nonword span should correlate more strongly with the discrimination test results than those of the identification or digit span tests.

A full correlation matrix using the data subset described above is shown in Table 5. The span test results in this analysis were only scored using the absolute span scoring method. The matrix shows that all of the measures were significantly intercorrelated. The only exception was the identification and digit span correlation. However, in this analysis, identification showed the strongest correlation with discrimination instead of nonword span or digit span. When short-term memory (digit span) was partialed out, the correlation coefficient for discrimination-identification dropped to only +0.6 ($p < 0.01$), while the discrimination-nonword span correlation coefficient was barely significant ($r = +0.39$, $p = 0.04$). When identification was partialed out, the correlation between discrimination and nonword span was no longer significant ($r = +0.32$, $p = 0.09$). In this analysis using a small subset of native nonwords from one talker, phonological coding played a much larger role in the discrimination of native nonsense words than a short-term memory capacity influenced by prior linguistic experience. If these results generalize to other samples of native nonsense words, they indicate that phonological coding, and thus prior linguistic experience, may play a greater role in the identification of more native-like than less native-like nonwords. With the less familiar phonemes and phoneme sequences of non-native nonwords, pure short-term capacity may begin to interact with long-term memory representations in the encoding of phonetic detail.

Test	**AXB**	**ID**	**DigABS**	**NonABS**	**WN**
AXB	X				
ID	0.66**	X			
DigAbs	0.44*	0.36	X		
NonABS	0.49**	0.42*	0.38*	X	
WN	0.43*	0.54**	0.51**	0.46**	X

Table 5. Correlations between the discrimination (AXB), identification (ID), digit span (DigABS), nonword span (NonABS), and the word-nonword (WN) condition of the paired-associate word-learning test. In these correlations, only the discrimination and identification results for the native nonsense word were used.
* $p < 0.05$, ** $p < 0.01$

Conclusions

The present study examined the contribution of phonological coding and short-term memory on the perception of non-native contrasts by native speakers of American English. The results of five speech perception and memory span tests demonstrated that short-term memory span correlated with the identification, but not discrimination, of non-native contrasts. However, a memory span task incorporating information in long-term memory concerning phonological similarity was shown to be the strongest predictor of perceptual performance in the discrimination test. The results support a model of short-term memory in which traces in short-term memory are augmented or transformed by information in long-term memory (Baddeley et al., 1998; Gathercole, Frankish, Pickering, & Peaker, 1999; Schweickert, 1993; Thorn & Gathercole, 1999).

These finding also support a model of cross-language speech perception that incorporates a short-term memory capacity regulating the extent of phonetic detail that is encoded in LTM. Currently, effects of short-term memory are not accounted for in either the Perceptual Assimilation Model or the Native Language Magnet model. In both of these models, the preservation of phonetic detail in the encoding of a non-native word in long-term memory is the consequence of their similarity to one or more native perceptual categories. To account for the results of this study, existing cross-language models should include a fixed short-term memory capacity that filters incoming speech prior to the identification process. Short-term memory could be modeled as a buffer in which phonological information is held prior to encoding in long-term memory, or perhaps as limited attentional resources for activating representations in long-term memory.

To strengthen the claim that short-term memory capacity is an important source of individual differences in cross-language speech perception, additional studies are needed correlating various measures of memory span and speech perception using a variety of listener groups and stimulus sets. In this study, only a small set of nasal consonants from two speakers of Malayalam was presented to monolingual speakers of English. Clearly, a greater range of stimuli (non-native vowel as well as consonant contrasts) and listener groups (multilingual, as well as monolingual speakers of languages other than English) should be used in future studies. In addition, the phonological similarity between stimulus materials in the span tasks and those in the speech perception tasks should be manipulated in several steps to examine the different roles that short-term and long-term memory may play in the encoding of non-native sounds that vary in their similarity to native sounds (roles alluded to in the analysis represented in Table 5). Hopefully, by focusing on individual differences and by examining the role of short-term on speech processing, we can gain a greater understanding of how non-native speech is perceived, encoded, and, ultimately, acquired by the learner.

References

Baddeley, A.D. (1986). *Working memory*. Oxford, England UK: Oxford University Press.

Baddeley, A.D., Gathercole, S.E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105(1)*, 158-173.

Baddeley, A.D., Papagno, C., & Vallar, G. (1988). When long-term learning depends on short-term storage. *Journal of Memory and Language*, *27*, 586-595.

Best, C. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171-203). Baltimore: York Press, Inc.

Bohn, O.-S. & Flege, J.E. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics*, *11*, 303-328.

Carpenter, P.A., Miyake, A., & Just, M.A. (1994). Working memory constraints in comprehension: Evidence from individual differences, aphasia, and aging. In M.A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 1075-1122). San Diego, CA: Academic Press.

Carroll, J.B. (1962). The prediction of success in intensive foreign language training. In R. Glaser (Ed.), *Training research and education* (pp. 87-136). Pittsburgh, PA: University of Pittsburgh Press.

Carroll, J.B. (1981). Twenty-five years of research on foreign language aptitude. In K. Diller (Ed.), *Individual differences and universals in language-learning aptitude* (pp. 83-118). Rowley, MA: Newbury House.

Flege, J.E. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, *15*, 47-65.

Flege, J.E. (1991). Age of learning affects the authenticity of voice onset time (VOT) in stop consonants produced in a second language. *Journal of the Acoustical Society of America*, *89*, 395-411.

Flege, J.E. (1992). Speech learning in a second language. In C. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, and application* (pp. 565-604). Timonium, MD: York Press.

Flege, J.E., Yeni-Komshian, G.H., & Liu, S. (1999). Age constraints on second language acquisition. *Journal of Memory and Language*, *41*, 78-104.

Gathercole, S.E. (1995). Is nonword repetition a test of phonological memory or long-term knowledge? It all depends on the nonwords. *Memory and Cognition*, *23*, 83-94.

Gathercole, S.E., & Baddeley, A.D. (1990). The role of phonological working memory in vocabulary acquisition: A study of young children learning new names. *British Journal of Psychology*, *81*, 439-454.

Gathercole, S.E., Frankish, C.R., Pickering, S.J., & Peaker, S. (1999). Phonotactic influences on short-term memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *25*, 84-95.

Gathercole, S.E., Service, E., Hitch, G.J., Adams, A.M., & Martin, A.J. (1999). Phonological Short-term memory and vocabulary development: Further evidence on the nature of the relationship. *Applied Cognitive Psychology*, *13*, 65-77.

Gathercole, S.E., Willis, C., Emslie, H., & Baddeley, A.D. (1991). The influences of number of syllables and wordlikeness on children's repetition of nonwords. *Applied Psycholinguistics*, *12*, 349-367.

Goldinger, S.D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, *105*, 251-279.

Harnsberger, J.D. (1998). *The perception of non-native nasal contrasts: A cross-linguistic perspective*. Doctoral dissertation, University of Michigan, Ann Arbor.

Harnsberger, J.D. (1999). A comparison of three metrics of perceptual similarity in cross-language speech perception. In S. Chang, L. Liaw, & J. Ruppenhofer (Eds.), *Proceedings of the Twenty-fifth Annual Meeting of the Berkeley Linguistics Society* (pp. 157-168). Berkeley, CA: Sheridan Books.

Harnsberger, J.D. (2000). A cross-language study of the identification of non-native nasal consonants varying in place of articulation. *Journal of the Acoustical Society of America*, *108*, 764-783.

Kuhl, P.K. & Iverson, P. (1995). Linguistic experience and the "perceptual magnet effect." In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 121-154). Baltimore: York Press, Inc.

Liberman, A.M., Harris, K.S., Kinney, J.A., & Lane, H. (1961). The discrimination of relative onset-time of the components of certain speech and nonspeech patterns. *Journal of Experimental Psychology*, *61*, 379-388.

Lively, S.E., Logan, J.S., & Pisoni, D.B. (1993). 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.

Lively, S.E., Pisoni, D.B., Yamada, R.A., Tohkura, Y., & Yamada, T. (1994). Training Japanese listeners to identify English /r/ and /l/ III: Long-term retention of new phonetic categories. *Journal of the Acoustical Society of America*, *96*, 2076-2087.

Logan, J.S., Lively, S.E., & Pisoni, D.B. (1991). Training Japanese listeners to identify English /r/ and /l/: A first report. *Journal of the Acoustical Society of America*, *89*, 874-886.

MacKain, K.S., Best, C.T., & Strange, W. (1981). Categorical perception of English /r/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics*, *2*, 369-390.

MacKay, I.R.A., Meador, D., & Flege, J.E. (2000). The identification of English consonants by native speakers of Italian. *Phonetica*, *58*, 103-125.

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

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A.M., Jenkins, J.J., & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of /r/ and /l/ by native speakers of Japanese and English. *Perception and Psychophysics*, *18*, 331-340.

Papagno, C., Valentine, T., & Baddeley, A. (1991). Phonological short-term memory and foreign language vocabulary learning. *Journal of Memory and Language*, *30*, 331-347.

Papagno, C., & Vallar, G. (1992). Phonological short-term memory and the learning of novel words: The effect of phonological similarity and item length. *The Quarterly Journal of Experimental Psychology*, *44A*, 47-67.

Papagno, C., & Vallar, G. (1995a). To learn or not to learn: Vocabulary in foreign languages and the problem with phonological memory. In R. Campbell & M. A. Conway (Eds.), *Broken memories: Case studies in memory impairment* (pp. 334-343). Oxford, England UK: Blackwell Publishers, Inc.

Papagno, C., & Vallar, G. (1995b). Verbal Short-term Memory and Vocabulary Learning in Polyglots. *The Quarterly Journal of Experimental Psychology*, *48A*, 98-107.

Schmidt, A.M. (1996). Cross-language identification of consonants. Part I. Korean perception of English. *Journal of the Acoustical Society of America*, *99*, 3201-3211.

Schweickert, R. (1993). A multinomial processing tree model for degradation and redintegration in immediate recall. *Memory and Cognition, 21*, 168-175.

Service, E. (1992). Phonology, working memory, and foreign-language learning. *The Quarterly Journal of Experimental Psychology*, *45A*, 21-50.

Service, E. (1998). The effect of word length on immediate serial recall depends on phonological complexity, not articulatory duration. *The Quarterly Journal of Experimental Psychology*, *51A*, 283-304.

Skehan, P. (1989). *Individual differences in second language learning*. London: Edward Arnold.

Strange, W. (1995). Cross-language studies of speech perception: A historical review. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 3-45). Baltimore: York Press, Inc.

Thorn, A.S.C., & Gathercole, S.E. (1999). Language-specific knowledge and short-term memory in bilingual and non-bilingual children. *The Quarterly Journal of Experimental Psychology*, *52A*, 303-324.

Trojano, L., & Grossi, D. (1995). Phonological and lexical coding in verbal short-term memory and learning. *Brain and Cognition*, *21*, 336-354.

Vitevitch, M.S., Luce, P.A., Charles-Luce, J., & Kemmerer, D. (1997). Phonotactics and syllable stress: Implications for the processing of spoken nonsense words. *Language and Speech*, *40*: 47-62.

Appendix

Word-Word pairs for the Paired Associate Word-Learning Task

- restaurant-skeleton
- finger-sheriff
- canyon-pepper
- tornado-computer
- tower-razor
- arena-family
- hamburger-passenger
- staple-neighbor

Word-Nonword pairs for the Paired Associate Word-Learning Task

- explosion-(pekʌrmʌn)
- actor-(fʊltaɪs)
- college-(pʌrnhʌs)
- leather-(sigmeb)
- manager-(remʌrnes)
- table-(heysʌk)
- physician-(kʌrndʌsmaɪd)
- telephone-(sʌtʌrsʌl)

This page left blank intentionally.