EFFECT OF LEXICAL ACCESS AND MEANINGFUL LINGUISTIC CONTEXT ON SECOND LANGUAGE SPEECH PERCEPTION

by

Albert O Jarvi

A thesis submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Master of Arts

Department of Linguistics

The University of Utah

December 2008

THE UNIVERSITY OF UTAH GRADUATE SCHOOL

SUPERVISORY COMMITTEE APPROVAL

of a thesis submitted by

Albert O Jarvi

This thesis has been read by each member of the following supervisory committee and by majority vote has been found to be satisfactory.

Chair: Rachel Hayes-Harb

THE UNIVERSITY OF UTAH GRADUATE SCHOOL

FINAL READING APPROVAL

To the Graduate Council of the University of Utah:

I have read the thesis of Albert O Jarvi in its final form and have found that (1) its format, citations, and bibliographic style are consistent and acceptable; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the supervisory committee and is ready for submission to The Graduate School.

Date

Rachel Hayes-Harb
Chair: Supervisory Committee

Approved for the Major Department

Chair/Dean

Approved for the Graduate Council

David S. Chapman
Dean of The Graduate School

ABSTRACT

Novel second language (L2) phonemic contrasts are difficult for learners to perceive and produce. Yet research has shown that even difficult L2 phonemic contrasts can be learned under some circumstances. Most of this research has been done in laboratory settings, using tasks that do not closely resemble natural communication. Among other characteristics, natural communication differs from these laboratory tasks in that (1) lexical access is usually not required in the laboratory tasks, and (2) target words in laboratory tasks are usually not embedded in meaningful linguistic context. This thesis describes an experiment designed to test whether these two characteristics of natural communication make it more difficult for learners to perceive L2 phonemic contrasts. The results of this experiment show that both requiring lexical access and embedding target words in meaningful linguistic context reduce the L2 learners' ability to use phonemic contrasts to distinguish L2 minimal pairs, but that processing meaningful linguistic context affects L2 perception more than does lexical access.

TABLE OF CONTENTS

LIST OF FIGURES

Figures

LIST OF TABLES

Tables

INTRODUCTION

Languages differ from one another in their systems of phonemic contrasts, and these differences between languages can create difficulties for second language (L2) learners. For example, English has a phonemic contrast between [l] and [ɹ], while Japanese does not have this contrast. Native Japanese speakers often perceive English [l] and [ɹ] as the Japanese phoneme /r/ (Aoyama, Flege, Guion, Akahane-Yamada & Yamada 2004). While these contrasts are difficult, research has shown that second language learners can exhibit improved perception and production of a novel L2 contrast under certain kinds of training conditions (see Curtin, Goad & Pater 1998). However, these laboratory-based studies may be limited in their ability to reflect actual second language acquisition and may not provide information about learners' ability to use the novel contrasts in more authentic situations. Other research has focused on L2 speech perception and production in naturalistic learning situations (e.g. Flege 1984; Flege, Takagi & Mann 1995). These studies recruit subjects who are living in L2 environments and test whether experienced L2 learners are more like inexperienced L2 learners or native speakers. These studies indicate that even with naturalistic experience, L2 learners neutralize nonnative contrasts in perception. For example, Takagi (2002) conducted an extensive experiment with a pretest, three phases of training (about six days each), and a posttest to test the limits of training Japanese speakers in the English /ɹ/-/l/ contrast. The subjects had received written English training in high school but had no spoken English

training. The author reported that sensitivity to the contrast increased for each subject, but the increase was not uniform across subjects. Takagi concluded that intensive laboratory training and naturalistic exposure does not seem to lead to native-like perception of /ɹ/ and /l/. These results are similar to observations made by foreign language instructors. L2 teachers often observe that the more 'authentic' the task, the more accented the learners' speech becomes. Foreign language teachers also observe that L2 learners can perform well in controlled classroom settings but revert to more accented speech in less controlled (i.e. more authentic) settings (see Hayes-Harb 2007).

It is possible that more 'authentic' tasks are more difficult for learners because (1) they require learners to access their lexical representation of L2 words, which may or may not accurately encode novel contrasts (Hayes-Harb & Masuda 2008), and/or (2) the task demands associated with authentic communicative tasks (i.e. operating in real time, incorporating background knowledge, perceiving novel phonemic contrasts in running speech, etc.) may demand the use of cognitive resources that are not then available to devote to the accurate perception of novel contrasts. Some studies which exhibit how requiring lexical access and increasing cognitive load by embedding target words in meaningful linguistic context affects L2 listening performance will be discussed later.

The purpose of this thesis is to document how well second language learners are able to perceive phonemic contrasts in order to differentiate second language words under varying task demands. The phonemic contrasts used in this thesis are palatalized and nonpalatalized dental consonants in Ukrainian. The task demands in this thesis will vary depending on whether or not the tasks require lexical access and whether or not the target words are embedded in meaningful linguistic context.

BACKGROUND

Perception and Production of Second Language Phonemes

Many theories exist explaining why L2 learners exhibit foreign accents. For example, foreign accents could result from the loss of basic speech learning mechanisms, inadequate phonetic input, or difficulty in preventing native language (L1) and second language (L2) systems from interacting (Flege, Takagi & Mann 1995). Many studies have focused on how L1 and L2 systems interact, specifically how L1 and L2 sound segments, which comprise phonemic contrasts, interact. A phonemic contrast consists of 'segment-sized constellations of phonetic properties that have become linguistically distinctive because they are used systematically to convey differences in word meanings' (Best 1994:169). Languages vary with regard to their phonemic contrasts. The L1 might have a phonemic contrast that the L2 does not have, or both the L1 and L2 might have the same phonemic contrast but use the phonetic information differently in categorizing phonemes. For example, along the voicing continuum, English has a two-way contrast between /b/ and /p/ while Thai has a three-way contrast between /b/, /p/, and /p^h/. These differences are important because they affect the way L2 sounds interact with L1 sounds when a second language is acquired. A large literature explores the difficulties that adult learners have acquiring the phonemic contrasts of a second language. Before this literature is discussed in detail, it is necessary to describe two of the more prominent models attempting to account for these difficulties; the Speech Learning Model (SLM;

Flege 1995) and the Perceptual Assimilation Model (PAM; Best 1994). Both the SLM and PAM predict that discriminability of phonemic contrasts is based on the relationship between L1 and L2 sounds. The main difference between the two models is that PAM is based on the ability to perceive articulatory properties and SLM is based on L2 learning (Guion, Flege & Loftin 2000).

Perceptual Assimilation Model

Infants are born with the ability to perceive most naturally occurring phonemic contrasts. Research has shown that after six months of age these infants start to become desensitized to contrasts that are not part of the phonological system of their native language. This is possibly due to experience with their native language. Presumably, children older than six months are less able to distinguish L2 phonetic contrasts because they start to develop L1 phonemic categories. Best's theory makes use of the fact that there seems to be a developmental change in speech perception. Young infant perception is different from older infant perception, which is different from adult perception. Using an ecological theory for speech perception, Best (1994) argues that initially children perceive linguistic sounds acoustically and can perceive contrasts in the same way adults can perceive a difference between a nonlinguistic 'snap' versus a 'beep'. At some point children recognize that linguistic sounds can be used for communication. Best argues that at this point, children begin to perceive articulatory gestures and categorize these gestures into phonemic information. This is the point where L1 begins to constrain the phonological system. Children begin to be less able to perceive nonnative contrasts. However, these constraints are not absolute; they can change over time.

Speech Learning Model

While age is an important factor in the speech learning model for L1 and L2 speech acquisition, the SLM focuses on L2 learning. For example, much of the previous research has shown that Japanese speakers seem to assimilate English /l/ and English /ɹ/ to Japanese /r/. The previous research also shows that /ɹ/ is less similar to Japanese /r/ than English /l/. One prediction of the SLM is that an L2 learner will learn to discriminate a phonemic contrast easier if the L2 phonetic segments are distinct from the L1 phonetic segments. Using the SLM, Aoyama et al. (2004) predicted that /ɹ/ will be perceived better than /l/ and that the earlier the L2 is learned the more likely new phonetic categories will be formed. Aoyama et al. (2004) tested this prediction by administering a discrimination test and a production test to children and adults at two different times. They reported that native Japanese (NJ) adults performed better during the first perception test than NJ children. But NJ children performed better at the second test than NJ adults. The authors concluded that for the children, perceptual learning took place for both contrasts /ɹ/-/l/ and /w/-/ɹ/. In the production experiment the NJ children improved for /ɹ/ and /w/ but not /l/, while NJ adults did not improve for any segment. The authors concluded that NJ children show more learning for /ɹ/ than /l/ and that the results provided support for the SLM hypothesis.

According to the SLM, adult learners who have received sufficient native-speaker input can master new L2 sounds, but L2 sounds which are classified as similar are more difficult. Many studies show that there is a difference between inexperienced L2 speakers, experienced L2 speakers, and native speakers. One such study was conducted by Flege, Munro & Skelton (1992). They show that nonnative speakers of English who

had been living in the United States performed worse on a production task than native speakers. In addition, more experienced L2 speakers performed no different than inexperienced L2 speakers. The authors suggested that this evidence supports the SLM hypothesis. These results support the SLM because the difference between inexperienced L2 speakers and experienced L2 speakers shows that L2 learners can improve their ability to perceive the novel contrast. These results also support the SLM because the difference between L2 speakers and native speakers shows that even though L2 speakers can improve their perception of novel contrast, it is difficult to reach native-like perception.

There are many hypotheses associated with the SLM. One of the most studied hypotheses is that an L2 learner will learn to discriminate a phonemic contrast easier if the L2 phonetic segments are distinct from the L1 phonetic segments. One difficulty in perceiving L2 phonemic contrasts is when two L2 phonemes are mapped to one L1 phoneme. This necessitates that the two native allophones be split into two L2 phonemes. The SLM predicts that an L2 speaker will have a more difficult time learning the contrast if the L2 phonemic contrast is an allophonic contrast in L1.

L2 Speech Perception and Production Studies

Despite the difficulties that learners exhibit with the perception and production of novel phonemic contrasts, it is clear that learners can and often do improve their ability to perceive and produce novel contrasts (Strange & Dittmann 1985; Logan, Lively & Pisoni 1991; Flege, Takagi & Mann 1995). The following sections discuss two types of studies that document this improvement: studies of L2 learners with naturalistic exposure to the

second language and laboratory training studies.

Naturalistic Studies

MacKain, Best and Strange (1981) show that Japanese L2 learners' ability to perceive English /ɹ/ and /l/ may evolve during naturalistic acquisition, while laboratory training is long and slow. Flege (1984) followed up on this research by conducting a naturalistic study. Flege compared the perception of English /s/ vs. /z/ in two groups of Arabic speakers and a native English group. An auditory word-picture matching test was used. The subject heard a word and then had to mark on the answer sheet which target word they heard (e.g. 'peas' or 'piece'). The answer sheet had two columns. On the top of the column was a picture corresponding to 'peas' or 'piece'. The inexperienced group had been in the United States less than two months. The experienced group had been in the United States on average 5.8 years (2.8-18). The analysis showed that the inexperienced group performed worse than the experienced group and the native group. The author concluded that hearing and speaking L2 in naturalistic conditions is better than training in laboratory conditions.

Flege, Takagi, and Mann (1995) is an example of a study where the subjects gained L2 experience by living in an L2 language environment. They studied phonemic contrasts in the context of Japanese speakers' production of English /ɹ/ and /l/. In this speech production experiment, three groups produced English speech tokens in three different speaking styles which the authors describe as definition, reading, and spontaneous. The first group included native English speakers (NE). The second group included native Japanese speakers who had lived in the United States at least 12 years

(EJ). The third group included native Japanese speakers who had lived in the United States for less than 3 years (IJ). The result of this experiment was that Japanese speakers that were experienced English speakers (EJ) produced the /ɹ/-/l/ contrast better than Japanese speakers that were inexperienced English speakers (IJ). The productions of the EJ group were similar to the NE group but their productions were not quite native-like. The authors concluded that even though it is extremely difficult for Japanese speakers to produce native-like contrasts, they were able to improve their ability to produce these contrasts.

The nature of L2 experience is an important issue in naturalistic studies. In a review of the literature on speech perception, Piske, MacKay, and Flege (2001) examined the methodology used in speech perception research. They examined the subject populations, the elicitation techniques, and rating techniques. In addition, Piske et al. (2001) also reviewed the factors which affect L2 speech acquisition; namely: age of arrival (AOA), length of residence (LOR), gender, formal instruction, motivation, and language use. The following studies have examined these factors in depth.

Piske et al. (2001) examined the effect of L1 use of foreign accent in early and late bilinguals. The subjects were native Italian speakers who had lived in Canada for various amounts of time. The authors concluded that native speakers of Italian who continue to speak their L1 have stronger foreign accents in L2 regardless of whether they were early or late bilinguals. They also concluded that age of L2 learning had a greater effect than L1 use.

Guion, Flege, and Loftin (2000) also used a population that learned the L2 in a naturalistic setting. Their purpose was to study the effect of L1 use on L2 production. The

authors concluded from this study that the L1 and L2 phonetic systems interact. The main effect that they found was that greater L1 use produced more accented L2 sentences, but L2 use did not affect the accent of L1 sentences.

Flege, Frieda, Walley, and Randazza (1998) studied the production of English /t/ by Spanish speakers. The Spanish speakers had learned English by living in the United States. They were divided into an Early Exposure and a Late Exposure group according to their age when they arrived in the United States. The authors investigated which factors affected the production of a nonnative phonemic contrast. Their results show that vowel height and number of syllables affected the voice onset time (VOT) values for English /t/ but that the lexical factors such as frequency, familiarity, cognate status, age of acquisition, and imageability did not.

Flege and MacKay (2001) were interested in whether there is a critical period to learning a second language and specifically learning a novel phonemic contrast. In this study native Italian speakers were compared with native English speakers both of whom lived in Canada. There were several groups of native Italian speakers based on the length of their residence in Canada. An oddity discrimination task and a classification task were used to test the subjects' ability to perceive English vowels. The authors concluded that the main effect in perceiving the contrast was AOA and L1 use. They also concluded that the ability to establish new vowel categories remains intact across the life span.

Flege and Liu (2001) investigated the effect of language experience in L2 learning. In Flege and Liu (2001) native Chinese speakers were divided by length of residence and by whether or not they were students. They tested the perception of English stops (/b d g p t k/). Two auditory grammar tests were given as well. The authors

concluded that LOR did not make a big difference between the groups. The biggest difference was between students and nonstudents. Presumably the students received more L2 input than nonstudents.

These studies recruited subjects that had learned the L2 on their own, and did not involve laboratory training. Without the controlled setting of the laboratory, it is difficult to measure how much L2 exposure the subjects received. Factors such as age of acquisition, age of arrival, length of residence, or L1 use have been suggested as measures of L2 exposure. One of the main difficulties is that these factors are all interrelated. To control for these and other factors, many studies have been conducted in laboratory settings. The next section describes a few of these studies.

Laboratory Training Studies

Many speech perception studies attempt to teach new L2 phonemic contrasts to speakers with no background in the L2. The subjects learn novel L2 phonemic contrasts by going through training sessions in the laboratory. Some of these studies have been successful at teaching novel phonemic contrasts. While these studies show that L2 learning is possible, it is important to note that these studies are limited in their ability to replicate natural speech. Specifically, these studies often lack the important factors of lexical access and meaningful linguistic context. The manner in which these studies lack the requirement of lexical access and meaningful linguistic context will be discussed after describing a few important studies.

An example of a standard perceptual study was done by McClaskey, Pisoni, and Carrell (1983). They extended previous work on VOT by using an auditory identification

task. In this type of study, subjects hear an auditory stimulus and must identify the word or sound. They attempted to generalize the perceptual training of VOT from one place of articulation to other places of articulation. They reported that subjects were able to identify three distinct categories along the voicing continuum and that this was transferred to other places of articulation. They concluded that new categories can be created in the laboratory.

An example of a standard perceptual training study was conducted by Bradlow, Pisoni, Akahane-Yamada, and Tohkura (1997). They investigated how perceptual training would affect native Japanese speakers' production of English /ɹ/ vs. /l/. None of the Japanese speakers had naturalistic exposure to English. The experiment consisted of three stages: a pretest, a training phase, and a posttest. The perception pretest consisted of a minimal pair identification task. In a minimal pair identification task, the subject sees two written words and hears one word. The subject must identify which word of the minimal pair matched the auditory stimulus. The perception posttest consisted of two generalization tasks. Bradlow et al. (1997) reported that /ɹ/ was identified more accurately than /l/ as a result of training. In addition, they reported that their results clearly demonstrate significant improvements in the Japanese trainees' productions of /ɹ/ and /l/ as a result of perceptual training. This training generalized to new words as well. The authors concluded that these results support the PAM which states that a unified common mental representation underlies both speech perception and speech production.

Heeren (2004) notes that early laboratory studies were not successful in training phoneme perception, while more recent studies have been more successful. One aspect of these studies is that written labels were given. Heeren (2004) forced the listeners to figure

out the acoustic differences themselves by presenting pictures. Native Dutch speakers with no English experience were tested on their perception of English /θ/ and /s/ by a four-interval AX discrimination task and an identification task before and after classification training. The training and identification tasks had pictures associated with the auditory input. The author reported that phoneme boundaries became closer to the English norm as a result of training.

Hayes-Harb (2007) examined the relationship between lexical and statistical information when learning second language phonemes. Hayes-Harb set up the training so that statistical information could be compared to lexical information. In the statistical condition, subjects were required to learn using the distribution of tokens. In the lexical condition, subjects were required to learn using pictures that were shown simultaneously with auditory stimuli. Hayes-Harb (2007) reported that both statistical and lexical information affected discrimination but that lexical information had a greater influence. In a second experiment, Hayes-Harb tested whether subjects would create lexical representations for new words that encoded a phonemic contrast. The author reported that lexical representations encoding the novel phoneme contrast were not created, and concluded from these experiments that there could be an intermediate stage of perceptual learning where speakers have a perceptual sensitivity to a novel contrast but are unable to represent the contrast lexically.

Curtin et al. (1998) conducted a study about whether contrastive aspiration or contrastive voicing is learned first. French and English speakers were trained on a Thai contrast between /b/-/p/-/p^h/. During the training, subjects heard a word and saw two pictures. The subjects had to choose which picture matched the auditory stimulus. The

test after the training was the same except the two pictures composed a minimal pair. An ABX discrimination task was also administered. The authors reported that French and English learners of Thai represent voice contrasts before aspiration contrasts. This is opposite of what previous research had shown. Pater (2003) attempted to replicate the Curtin et al. (1998) study. Pater changed the study so that the lexical task and the nonlexical task only differed in whether lexical access was required. The training in this study taught meanings of words through repeated sound-picture trials. Pater (2003) found the opposite of Curtin et al. (1998). Specifically, aspiration was discriminated better than voice in both tasks.

A few studies have incorporated both naturalistic exposure and laboratory training. Logan, Lively, and Pisoni (1991) reported on their attempts to train Japanese speakers to identify the /ɹ/-/l/ contrast. The Japanese subjects had lived in the United States but were not comfortable with their speaking ability. The study had a pretest, about three weeks of training, and a posttest. The authors reported that subjects were able to transfer what they learned about /ɹ/ and /l/ during training to the posttest stimuli. Hayes-Harb and Masuda (2008) also trained subjects who had naturalistic exposure to the L2. The purpose of the experiment was to investigate the ability to lexical encode contrasts at different stages of learning. They tested three groups of subjects on their discrimination of Japanese consonant length. The first group of native English speakers had no Japanese experience. The second group of native English speakers had one year of Japanese experience. The third group consisted of native Japanese speakers. The subjects were taught words by presenting auditory stimulus at the same time as pictures. The perception task consisted of an auditory word-picture matching task. For the production task, the

subjects had to name a picture. Performance of the native English speakers with one year of Japanese experience was between that of native Japanese and less experienced L2 speakers. As has been shown, the study of L2 speech perception can be done with subjects who learn the second language in naturalistic settings and in laboratory training sessions.

The Effect of Increasing Cognitive Load on Speech Perception Tasks

Another aspect of L2 speech perception studies involves the cognitive load of the tasks of the study. For example, some researchers have studied the effect of varying the tasks demands on L2 speech perception. Three such studies are described in detail below.

Bradlow and Pisoni (1999) investigated the influence of various factors on word recognition by native and nonnative listeners. They noted that factors such as a high degree of stimulus variability, difficult listening conditions, and lexical characteristics were important in speech perception. In one experiment, native English speakers and nonnative English speakers were tested on a word recognition test. Target words were separated into an easy word list and a hard word list based on word frequency, neighborhood density, and neighborhood frequency. The easy word list consisted of words that had a higher mean word frequency, a lower mean neighborhood density, and lower mean neighborhood frequency as compared to the hard word list. The authors reported that easy word lists were transcribed more accurately than hard word lists. They suggested that since a lexical competition effect is observed under highly favorable listening conditions in the laboratory, the effect would be even greater under less favorable conditions. In another word recognition experiment, the stimulus varied on

easy vs. hard and on whether there was a single talker or multiple talkers. In this experiment, the nonnative listeners displayed lower word recognition scores than the control group. However, the nonnative listeners displayed the same patterns as the control group in that they recognized more words in the easy word list compared to the hard word list and they recognized more words when presented with stimuli from a single talker as opposed to multiple talkers. The authors concluded that listeners tune and adjust their speech perception mechanisms to take advantage of surface level or paralinguistic consistencies in the signal such as signal-related, lexical, or instance-related factors.

Pater, Stager, and Werker (2004) conducted an experiment regarding word learning in children. The main issue centered on whether children have the ability to perceptually represent phonological structures and encode them in lexical representations. Research has shown that infants can perceive most phonological contrasts and that children older than one year begin to organize their perception of phonological contrasts (see Pater, Stager & Werker 2004). The authors explain the previous research on speech perception in children by claiming that when a contrast is first acquired it is not stable and can be lost under high processing demands. In their experiments, 14-month-old infants were engaged in a word learning task. In these experiments, the infant heard a word and saw a moving picture. This required the infants to pair a meaning distinction with a sound distinction. Earlier work by Stager and Werker (1997) and Werker, Fennell, Corcoran, and Stager (2002) showed that 14-month-old infants failed to respond to a switch between /bi/ and /di/ when engaged in this type of word learning task. These authors suggested that the diminished availability of resources for phonological processing results in the construction of a phonological parse that is reduced in

complexity. Pater, Stager, and Werker (2004) replicated these results when using more complex contrasts: /bin/ vs. /din/ and /bin/ vs. /p^hin/. Pater et al. (2004) suggested that contrasts, before being fully acquired, are partially integrated into the phonological system, during which time their maintenance is affected by processing demands, such as the establishment of sound and meaning pairings.

Another example of manipulating cognitive load comes from a speech production study. In conducting speech perception research, it is important to control the speech samples that are used. For example, Harnsberger, Wright, and Pisoni (2007) developed a new technique to elicit hypo-articulated, citation, and hyper-articulated English speech while controlling the linguistic content from English speakers. To elicit the hypo-articulated, or reduced, speech, the authors introduced a digit span task that increased the memory load of the speaker. The subject first saw a sequence of single digits (0-9) on a screen. Then the subject saw a sentence and was asked to read aloud the sentence. Finally, the subject was asked to recall the digit sequence in the order it was presented. For the citation speech, subjects were asked to read aloud a sentence they saw on the screen. For the hyper-articulated speech, subjects were asked to read aloud a sentence. A prompt, "Please read the sentence more clearly" was repeated twice. They reported that 80% of the subjects produced three distinct speaking styles. In their study they increased the cognitive load of the speaker to elicit reduced speech. If increasing the cognitive load of a speaker affects speech production, it is likely that increasing the cognitive load of a listener by requiring subjects to access their lexicon and to process meaningful linguistic context affects speech perception as well.

Bridging the Gap: Making Laboratory Tasks More 'Naturalistic'

As mentioned earlier, studies conducted in laboratory settings typically differ importantly in at least two ways from natural communication. First, they often do not require lexical access. For example, reading from word lists does not require lexical access because words are written. A listener can gather phonemic information from the orthography rather than accessing the lexicon. Seeing the orthographic representation of the word is different from seeing a picture of an object and accessing the lexicon to determine the word to which the picture corresponds.[1] Second, in laboratory tasks, auditory stimuli are typically presented in isolation (i.e. not in running speech). In the rare cases where auditory stimuli are presented in running speech, typically a carrier sentence is used which does not provide meaningful linguistic context (e.g. 'I like to say ___ more than ___.').

Table 1 lists some of the methods used by past speech perception studies, and Table 2 lists some of the methods used by past speech production studies. These tables show what tasks were done and whether lexical access was required and whether meaningful linguistic context was used. These tables show that some studies have required lexical access but only a few studies have embedded the target word in a meaningful linguistic context.

As shown in Table 1 and Table 2, researchers have developed and used many tasks to study speech perception and production. They have used tasks such as phoneme identification, discrimination, accentedness judgments, word recognition, and word

---

[1] On the other hand, it is important to note that lexical access is not required when phonographic orthography is used. It may be different if, for example, Chinese characters are used.

Table 1. A review of past studies of second language speech perception

Study	Task	Lexical access required?	Use of meaningful linguistic context required?
Pisoni & Lazarus 1974	random identification	no	no
	sequential identification	no	no
	ABX discrimination	no	no
	4IAX discrimination	no	no
McClaskey et al. 1983	identification	no	no
Flege 1984	auditory word-picture matching	yes	no
Strange & Dittman 1984	minimal pair	no	no
	identification	no	no
	oddity discrimination	no	no
	AX discrimination with feedback	no	no
Werker & Logan 1985	AX discrimination with ISI	no	no
Logan et al. 1991	identification	no	no
Flege et al. 1992	listening test	no	no
Flege et al. 1995	vocabulary test	yes	no
	forced-choice identification	no	no
	foreign accent task	no	no
Bradlow et al. 1997	identification	no	no
Curtin et al. 1998	auditory word-picture training	yes	no
	identification	yes	no
	minimal pair	yes	no
	ABX discrimination	no	no
Ingram & Park 1998	identification	no	no
	discrimination	no	no
Bradlow and Pisoni 1999	word recognition	no	no
	word familiarity	no	no
Walley & Flege 1999	identification	yes	no
Guion, Flege & Loftin 2000	foreign accent task	no	no
Guion, Flege, Akahane-Yamada & Pruitt 2000	identification	no	no
	goodness of fit to Japanese consonant	no	no
	categorical discrimination	no	no
Flege & Liu 2001	identification	no	no
	grammaticality judgment	no	yes
	LCT aural grammar test	no	yes

(continued on next page)

Table 1. (continued)

Study	Task	Lexical access required?	Use of meaningful linguistic context required?
Piske et al. 2001	foreign accent task	no	no
Takagi 2002	identification	no	no
Bent & Bradlow 2003	word recognition	no	yes
	word familiarity	no	yes
Pater 2003	auditory word-picture training	yes	no
	XAB discrimination	no	no
	XAB discrimination	yes	no
	categorical discrimination	no	no
Weber & Cutler 2004	word recognition	yes	yes
Aoyama et al. 2004	categorical discrimination	no	no
Cutler & Otake 2004	auditory lexical decision	no	no
Flege & MacKay 2004	oddity discrimination	no	no
	classification with rating	no	no
	vowel identification	no	yes
Heeren 2004	classification with feedback	yes	no
	4IAX discrimination	no	no
	identification	yes	no
Pater et al. 2004	word learning/fixation task	yes	no
Escudero & Boersma 2004	forced identification task	no	no
Harnsberger et al. 2007	discrimination	no	no
Hayes-Harb 2007	statistical training	no	no
	lexical training	yes	no
	discrimination	no	no
	word learning phase	yes	no
	matching test	yes	no
Hayes-Harb & Masuda 2008	auditory word-picture matching	yes	no

Table 2. A review of past studies of second language speech production.

Study	Task	Lexical access required?	Use of meaningful linguistic context required?
Flege et al. 1992	elicitation task	no	no
Flege et al. 1995	read definitions	yes	no
	read words in isolation	no	no
	create novel sentences using words from word list	yes	no
Bradlow et al. 1997	word list and auditory prompts	no	**no**
Flege et al. 1998	word knowledge test	no	no
	delayed repetition	no	no
Guion, Flege & Loftin 2000	repetition	no	no
Guion, Flege, Akahane-Yamada, & Pruitt 2000	repetition	no	**no**
Piske et al. 2001	delayed repetition	no	no
Aoyama et al. 2004	elicitation	yes	no
Harnsberger et al. 2007	digit span task	no	no
	reduced elicitation	no	yes
	citation elicitation	no	no
	hyper-articulated elicitation	no	no
Hayes-Harb & Masuda 2008	picture naming	yes	no

familiarity tasks. Auditory identification and word recognition tasks are similar in that subjects hear a word or sound segment and are required to identify the word or sound segment. The stimuli used are mainly in isolation, and, occasionally, written words are presented. There are many different types of discrimination tasks. In these tasks, the subject hears at least two different words or sound segments. In an AX task, the subject must decide if the X stimuli was the same as the A stimuli. In an ABX, AXB, or XAB task, the subject must decide whether the A stimuli or the B stimuli was the same as the X stimuli. These tasks may or may not provide written words and rarely provide pictures. These auditory stimuli are in isolation, in small phrases, or in carrier phrases. As

mentioned earlier, carrier phrases are set phrases in which the target stimuli are embedded; identification of the target words does not require that subjects accurately process this linguistic context. In accentedness judgment, subjects produce speech samples from word lists or auditory repetition tasks. Other subjects then judge the word or sound segment in question based on whether it was foreign sounding or not. In word familiarity tasks, subjects are asked to judge whether a word, which is usually presented in isolation or using orthography, is familiar or not. Without manipulating the tasks, these auditory speech perception tasks do not require lexical access or the processing of meaningful linguistic context.

In review, studies using word list tasks do not require lexical access or the use of meaningful linguistic context (Flege et al. 1992; Bradlow et al. 1997; Flege et al. 1998). Studies using auditory stimuli and pictures do require lexical access but do not use meaningful linguistic context (Curtin et al. 1998; Pater 2003; Hayes-Harb 2007). On the other hand, studies using carrier sentences do not require lexical access but do require using the context, even though it is not meaningful linguistic context (Flege 1984).

Two studies are worth mentioning regarding their study design. First, Flege and Liu (2001) made use of meaningful linguistic context by administering two grammar tests. These grammar tests required the subjects to use meaningful linguistic context to answer the questions. While these tasks are important, the tasks were not designed to test speech perception. Second, Weber and Cutler (2004) manipulated both lexical access and meaningful linguistic context to study L2 speech perception. The purpose of the study was to determine whether phonetically similar words in either L1 or in L2 compete against each other during a word-recognition task. A typical trial in the experiments

conducted by Weber and Cutler (2004) showed four black and white line drawings as well as four shapes in the corner. In each trial, an instruction was given such as: *Click on the panda. Now put it on top of the circle.* In this example, there would be a picture of a panda, a pencil, and two distracter pictures. Panda would be the target word, while pencil would be the competitor word. Lexical access was required because they had to decode the pictures. The target word was embedded in meaningful linguistic context. The results of their experiments showed that competitor words increased the fixation time on the target words for nonnative speakers but not for native speakers. The present thesis is different from their study because this thesis investigates the individual effects of both lexical access and the use of meaningful linguistic context on learners' performance on listening tasks, and involves a more complex contextual situation.

Native English Speakers Learning Ukrainian

The case of native English speakers learning Ukrainian provides an opportunity to study the effect of lexical access and meaningful linguistic context on perceiving a novel L2 phonemic contrast. The phonemic contrast studied in this experiment is the contrast between palatalized dental consonants and nonpalatalized dental consonants in Ukrainian.

Ukrainian

The contrast in palatalization in Ukrainian comes from a change which occurred in Old Russian, from which Ukrainian developed. Most consonants became 'paired' for the palatalization contrasts (Padgett 2003). In Ukrainian, this pairing is between a 'hard series' of nonpalatalized consonants and a 'soft series' of palatalized consonants. The 'hard series' of dental consonants consists of the following sounds: /t/, /d/, /s/, /z/, /ʦ/,

/ʒ/, /n/, /l/, and /r/. The corresponding palatalized dental consonants are /tʲ/, /dʲ/, /sʲ/, /zʲ/, /ʦʲ/, /ʒʲ/, /nʲ/, /lʲ/, and /rʲ/ respectively (Shevelov 1993).

It is important to discuss the orthographic representation of palatalization in Ukrainian. In the orthographic system of Ukrainian, there is one set of consonants and two sets of vowels. The orthographic representation of the consonants corresponds to the nonpalatalized set of consonants. The orthographic representation of the vowels is split into a 'hard series' and a 'soft series'. The 'soft series' is defined as the same as the 'hard series' of vowels except the 'soft series' has a /j/ sound preceding the vowel. Learners of Ukrainian learn to palatalize the consonant when the consonant is followed by a 'soft series' vowel. This obscures the fact that Ukrainian has six vowels (/i/, /e/, /y/, /a/, /o/, and /u/) and two sets of consonants (Shevelov 1993). For example, one minimal pair in Ukrainian is [rɨs], рис 'rice,' and [rʲis], pic 'grew'. Orthographically it seems that the contrast is in the vowel. The surface forms of these words seem to indicate that there are two contrasts (i.e. palatalization and vowel quality) and therefore these words are not minimal pairs. It is beyond the scope of this thesis to decide the underlying form of these contrasts, but it is assumed that the underlying contrast is between the consonants. This assumption is made because there is allophonic variation of the vowel. The vowel is always [i] after a palatalized consonant such as [rʲis] and [ɨ] after a nonpalatalized consonant such as [rɨs]. For this reason, the stimuli in this experiment did not contrast in word initial position before /i/ (for more detail regarding this problem see Rubach 2002; Rubach 2007; Padgett 2003). Another way to represent palatalization is by using the letter < ь >. This is called the soft sign. It is only used after consonants to indicate that the preceding consonant is 'soft' or palatalized. In order to reduce any effects caused by

vowels and maintain the orthographic conventions of Ukrainian, six out of the seven minimal pairs in this experiment used consonants followed by the soft sign.

English

It is difficult for native English speakers to perceive the palatalization. One possible explanation is described by Weber and Cutler (2004). They concluded that learners can encode a contrast lexically but do not perceive the contrast. One example of this would be if phonemic contrasts are represented lexically but that input processing only maps the input to one of the phonemes (i.e. /æ/ and /ɛ/ are both mapped to /ɛ/. This explanation is similar to the author's experience while learning Ukrainian. The author often felt that he knew that a contrast existed between words like /rɨs/ and /rʲis/, but that he could not perceive or produce the contrast. This experience was anecdotally supported by many of his colleagues who also learned Ukrainian.

Hypothesis

It is hypothesized that nonnative speakers will recognize novel phonetic contrasts better in a task that requires no lexical access and does not use meaningful linguistic context compared to a task which requires both lexical access and the use of meaningful linguistic context.

METHODS

Subjects

Twelve subjects were recruited for the experiment over a three month period. The subjects were living in the Salt Lake City, Provo, or Ogden areas in Utah at the time of recruitment. Eleven were males and one was female. They were all between 21 and 30 years of age. No subject disclosed that they had a hearing disorder or were taking any medications that might affect their motor skills. Their native language was English. All the subjects learned Ukrainian by serving as a missionary for their church. Each subject had two months of intensive language training before moving to Ukraine. Each subject lived in Ukraine for about 22 months, except for one subject who stayed for about 16 months. The subjects had completed their missionary service at different times within the last seven years. It is unknown how often each subject used Ukrainian after returning to the United States. Nine subjects felt they were fluent in Ukrainian, while two subjects felt they were conversational, and one subject had a basic level of understanding. Nine of the subjects participated in a quiet room, while three subjects participated in a sound-controlled booth. All subjects were paid $10 for participating. Two subjects were removed from the analysis due to experimenter error.

Stimuli

The target words for this experiment consisted of 14 Ukrainian nonwords of the form CVC. These 14 words consisted of seven minimal pairs. One set of seven contained a nonpalatalized consonant, while the other set of seven contained a palatalized consonant and formed minimal pairs with the first set (e.g. /fɬt/ фит and /fɬtʲ/ фить).[2] Nonwords were created for this experiment in order to control for subjects' previous vocabulary. Five of these words contrast in the word-final position.[3] Two of these words contrast in the word-initial position. As noted earlier, the orthographic representations differ in whether 'ь' is present or not. The vowels of the /luʃ/-/lʲuʃ/ луш-люш pair were selected according to Ukrainian orthographic conventions. These target nonwords and the pictures of the nonobjects that were associated with each target nonword are listed in Table 3. The pictures have been normalized so that on average each picture is equally unrecognizable to the subjects.

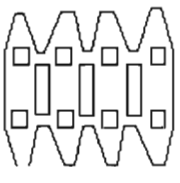
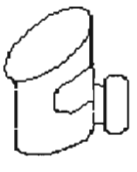
The presentation of the visual stimuli was different for each task. For Task A, written words in Ukrainian orthography were presented in the middle of the screen. For Task B, black and white line drawings of nonobjects were presented in the middle of the screen. For Task C and Task D, four objects were shown with one object in each corner

---

[2] Originally, two sets of nine target nonwords were created based on the nine phonemic contrasts in Ukrainian. Five of the minimal pairs contrasted in word final position, while four of the minimal pairs contrasted in word initial position. Later, it was necessary for two sets to be removed for at least two reasons. First, the number of target words was reduced from eighteen nonwords to fourteen nonwords in order to reduce the time it would take for subjects to complete the experiment. Second, in consultation with Speech Acquisition Lab meeting, the lab members felt that two sets were confusing and should be removed. These two sets contained the novel contrast in the word initial position. In addition, by removing these two sets of minimal pairs, two of the phonemic contrasts in Ukrainian were also removed.

[3] It is possible that due to word final devoicing only five phonemic contrasts are examined rather than seven contrasts. The subjects might have heard the sounds [d] and [z] as /t/ and /s/, respectively.

Table 3. List of target nonwords in Ukrainian

Contrast	Hard		Soft	
/t/-/tʲ/	фит /fɨt/		фить /fɨtʲ/	
/d/-/dʲ/	гуд /ɦud/		гудь /ɦudʲ/	
/s/-/sʲ/	кас /kas/		касьь /kasʲ/	
/z/-/zʲ/	боз /boz/		бозь /bozʲ/	
/ʦ/-/ʦʲ/	вец /veʦ/		вець /veʦʲ/	
/n/-/nʲ/	нох /nox/		ньох /nʲox/	
/l/-/lʲ/	луш /luʃ/		люш /lʲuʃ/	

of the screen. In these latter two tasks, the top two pictures corresponded to a minimal pair. The order of the minimal pair was randomized. For Task C, the top two objects were the written words in Ukrainian orthography that were used in Task A. For Task D, the top two objects were the black and white line drawings of nonobjects which where used in Task B. In both Task C and Task D, the bottom two objects were black and white line drawings of an airplane and a train. The order of the airplane and the train was randomized. Pictures of an airplane and a train were chosen because both the word /lʲɨtak/, літак 'airplane' and the word /pojiizd/, поїзд 'train' are two-syllable, masculine words in Ukrainian. These number of syllables and gender were selected in order to minimize the possibility that subjects would recognized the difference between syllables or gender rather than the perceptual difference between the minimal pair.

The auditory stimuli were recorded on a Marantz Professional Solid State PMD660 recorder in a sound-controlled booth during one session. The talker was an adult male native speaker of Ukrainian. The talker had recently moved from the city of Chernivtsi in western Ukraine to the United States. The target words were elicited by placing the target word at the end of a carrier phrase for the recording. The carrier phrase that was used is provided in example (1).

(1) Мені подобається сказати слово _____.
I-1st.dat like-3rd.sing say-inf word-nom _____.
I like to say the word _____.

There were a total of 14 target word sentences. Each sentence was recorded three times. Only the second recording was used. The target word was cut from the carrier phrase and normalized for peak-intensity. In Task A and Task B, these target words were presented in isolation. In Task C and Task D, these target words were embedded in a question.

Additional sentences were also recorded for the Task C and Task D which are described in detail in the next section. For Task C and Task D, the subjects were told that the color of the airplane was either yellow or blue and the color of the train was the opposite. The words /ʒovtɨj/, жовтий 'yellow' and /sɨnʲij/, синій 'blue' were chosen because in Ukrainian they are two syllable words. Two syllable words were chosen to minimize the possibility that subjects would recognize a syllable difference instead of the phonemic contrast. An example of this type of sentence is provided in example (2).

(2) Синій літак та жовтий поїзд.
blue-nom airplane-nom and yellow-nom train-nom.
The airplane is blue, and the train is yellow.

There were four possible arrangements for this sentence due to the two colors and two objects. Each of these four sentences was recorded three times. Only the second recording was used. Each recording was normalized for peak-intensity.

The main prompt for Task C and Task D is a question. The talker recorded two versions of this question; one in which the color is blue and one in which the color is yellow. The word 'table' (*stil*; стіл) was used as a filler word in place of the target word. Example (3) provides an example of this question.

(3) Чи стіл над тому, що синій?
is table-nom above that-loc, which blue-nom?
Is the table above that which is blue?

These two sentences were recorded three times and only the second recording was used. These two recordings were cut into three segments; the question particle /ʧɨ/ Чи, the filler word, and the rest of the question. The filler word was not used in the experiment. The other two segments were normalized for peak-intensity.

For each trial in Task C and Task D, subjects heard the auditory prompt in the

following order. First, subjects heard the sentence that described the color of the airplane and train. After a 2 second pause, subjects heard the question particle *chi*. After a 50 millisecond pause, subjects heard the target word. After another 50 millisecond pause, subjects heard the rest of the question. The pauses before and after the target word were added to increase the understandability of this auditory stimulus.

Procedures

The experiment was conducted using the DMDX software. All of the words and pictures displayed were in black and white. All of the pictures except for the airplane and train in Tasks C and D are nonsensical. Because of the limited number of subjects, the order of the tasks was randomized. Each subject was randomly assigned a different order of tasks, but due to experimenter error two subjects received the same order (subjects 9 and 10). The subjects sat at a computer and listened to the auditory stimuli using headphones. The subjects responded by pressing a button labeled 'Yes' or a button labeled 'No' on the keyboard. The subjects were instructed not to press any other key except the spacebar which allowed the subject to move from one task to another.

The experiment began by training the subject on the Ukrainian nonwords. This was done by showing a black and white line drawing of a nonobject on a computer screen. At the same time, the subject heard the nonword in isolation. The subject was required to connect the auditory stimulus with the visual picture. Each word was presented four times in a random order. This training was repeated after the second and third tasks except each word was only presented once. After the initial training, a test was given to verify that the subject had indeed memorized the new nonwords. During this test

the subject simultaneously saw a nonobject and heard a nonword. Each nonword was presented twice, once with the correct nonobject and once with an incorrect nonobject. In this training test, the two were not part of a minimal pair. There were a total of 28 trials in this test. The subject did not begin the experiment until they learned at least 90% of the nonwords.

The experiment consisted of four tasks. These four listening tasks were designed to indicate the subjects' ability to discriminate novel phonemic contrasts under different conditions. The tasks were manipulated according to whether or not they required lexical access and whether or not the target word was embedded in meaningful linguistic context. A brief description of these tasks is given in Table 4.

In the No-Lexical-Access conditions (A and C), lexical access was not required. By showing the orthographic representation of a word, the subject was given phonetic information regarding the phonemic contrast (i.e. the presence of a soft sign which represents whether the consonant is palatalized or not). While it was possible that the subjects would access their lexical representations of the words, the subject critically did not NEED TO access their lexical representation of the words because the subject had this

Table 4. Four tasks involving Lexical Access and/or Meaningful Linguistic Context

	No-Context	**Context**
No-Lexical-Access	A. A word reading task in isolation, with no lexical access required	C. A word reading task in meaningful linguistic context, with no lexical access required.
Lexical-Access	B. A picture naming task in isolation, where lexical access is required.	D. A picture naming task in meaningful linguistic context, where lexical access is required.

additional orthographic information. In the Lexical-Access conditions (B and D), subjects needed to access their lexicon in order to retrieve the phonological form of the target word. In the No-Context conditions (A and B), the target words were not embedded in meaningful linguistic context. In the Context conditions (C and D), subjects must both (1) make use of meaningful contextual information to answer appropriately (i.e. keep track of what is blue and what is yellow) and (2) perceive the target word in the context of running speech (i.e. not in isolation). The next four sections describe each task in detail.

Task A. No-Lexical-Access and No-Context

This task is similar to many of the previous studies. In this task, a written word is shown on the computer screen. At the same time, the subject hears a word which either (1) corresponds to the written word on the screen or (2) corresponds to the word which is a minimal pair with the written word on the screen. The difference between the words is palatalization as described above. The subject is instructed to determine whether the word which the subject heard is the same as the word on the screen by answering Yes or No. A Yes means that they are the same. A No means that they are different. By showing the written words, the subject does not need to access their lexicon in order to respond as they may simply decode the written forms. By listing only one word, there is no need to process meaningful linguistic context. Figure 1 is an example of the screen during Task A in which a target word is heard in isolation.

During this test the subject saw a written word and heard a target word. For example, in Figure 1, the target word /ɦudʲ/ is presented in its orthographic form. Each

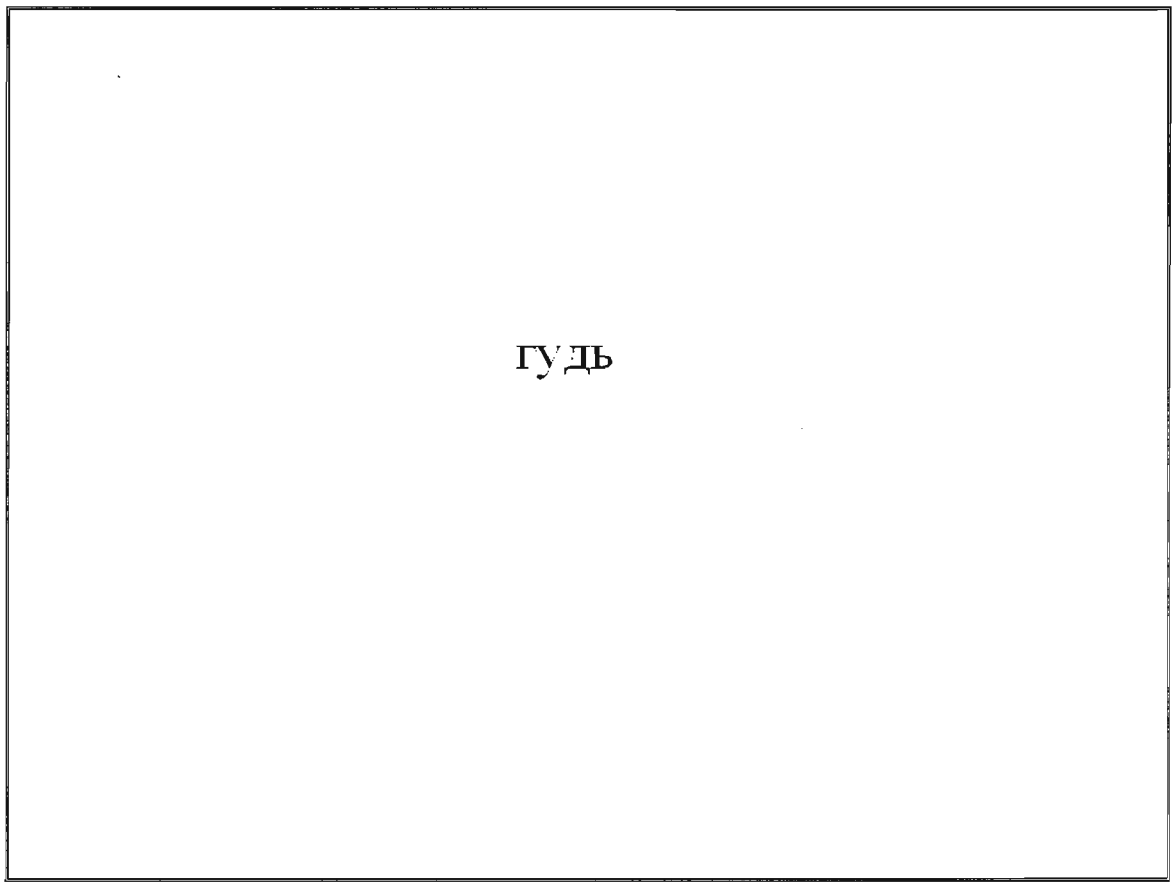


Figure 1. Screenshot of Task A, Target: /ɦud^{j}/

target word was presented twice, once where the target word was written and once where the corresponding minimal pair was written. There were 28 trials in this test.

Task B. Lexical-Access and No-Context

This task is similar to the nonlexical and noncontextual task. However, in this task, the subject sees a picture on the computer screen. At the same time, the subject hears a word that either (1) corresponds to the picture on the screen or (2) corresponds to the word that is a minimal pair with the picture on the screen. Again, the difference between the words is palatalization. The subject is instructed to determine whether the word that the subject heard is the same as the word that corresponds to the picture on the

screen by answering Yes or No. A Yes means that they are the same. A No means that they are different. The only difference between Task A and Task B is that during Task B the subject will see pictures instead of written words. By showing pictures, it is assumed that the subject will need to access their lexicon to retrieve the words which correspond to the picture. There is still no meaningful linguistic context involved because the auditory stimuli are words in isolation. Figure 2 is an example of the screen during Task B in which a target word is heard in isolation.

During this test the subject saw a nonobject and heard a target word. For example,

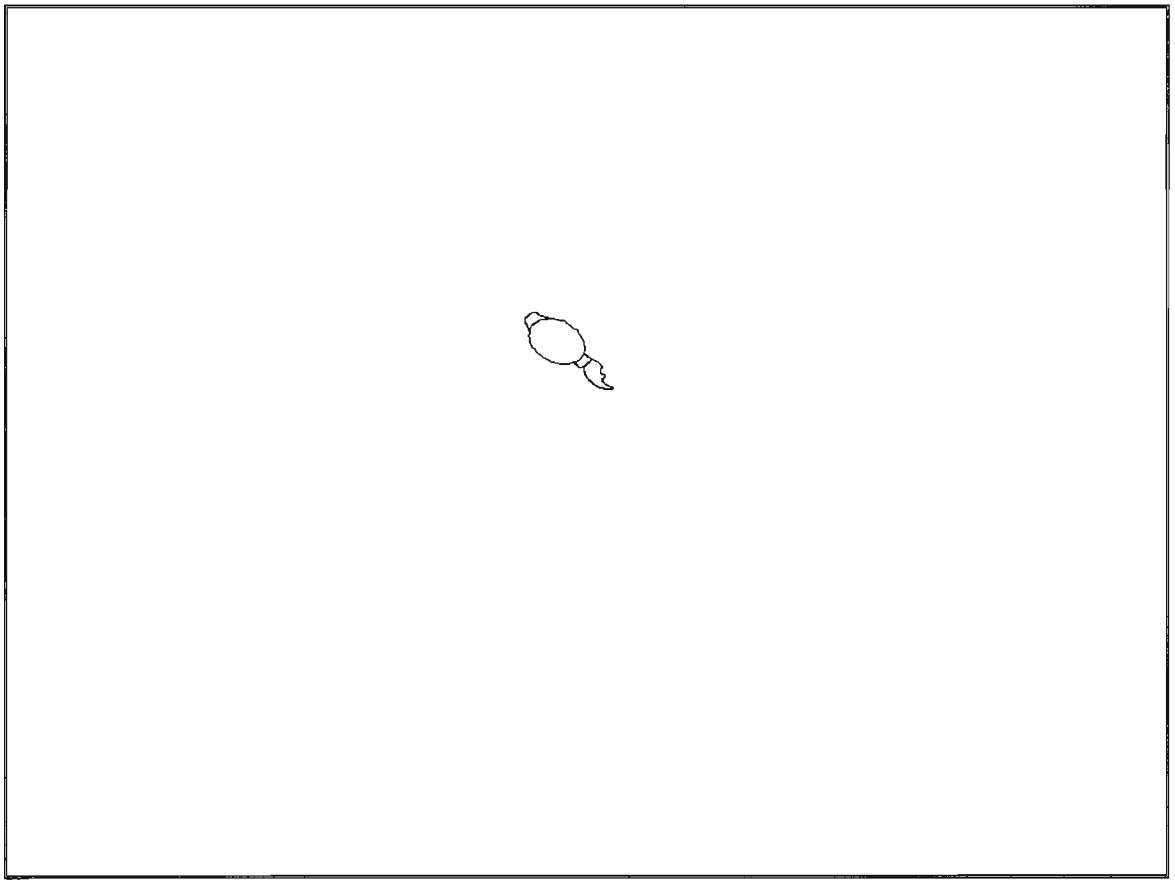


Figure 2. Screenshot of Task B; Target: /nox/

in Figure 2, the target word /nox/ is presented in its picture form. Each target word was presented twice, once where the nonobject corresponding to the target word was shown and once where the nonobject corresponding to the minimal pair of the target word was shown. There were 28 trials in this test.

Task C. No-Lexical-Access and Context

This task tests the subject's ability to distinguish phonemic contrasts in words which are embedded in a meaningful linguistic context. At the beginning of the task, the subject is instructed to answer Yes or No to the question which they will hear. In each trial, the subject will see two written words displayed over black and white line drawings of an airplane and a train. The two words will be a minimal pair differing in palatalization. One word will be above the airplane and the other word will be above the train. A prompt will inform the subject of the color of the airplane and the train. Then the target word is embedded in a question which asks whether the target word is above the object that is a certain color. The subject must remember the colors of the airplane and the train and discriminate between the minimal pair to correctly answer the question. The color and order of the airplane and train as well as the order of the words are randomized in each trial. By presenting the written word, the subject has no need to retrieve their lexical representation of the word. On the other hand, the subject is required to decipher the meaningful linguistic context in which the word is embedded to determine the correct answer. Example (4) provides an example of the auditory prompt the subject will hear while seeing a picture such as Figure 3 on the screen.

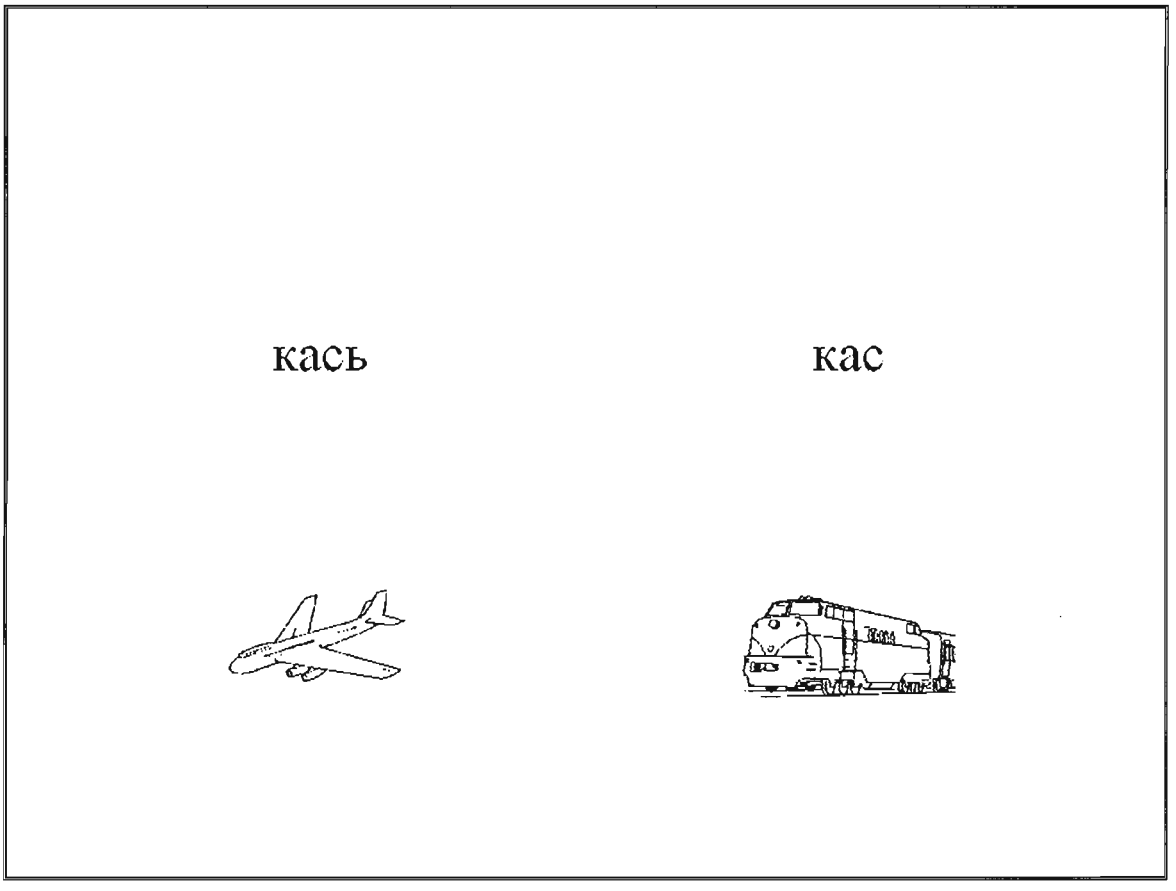


Figure 3. Screenshot of Task C; Target: /kas/

(4)	Синій	літак	та	жовтий	поїзд.	Чи
	blue-nom	airplane-nom	and	yellow-nom	train-nom.	Is
	кас	над	тому,	що	синій?	
	/kas/-nom	above	that-loc,	which	blue-nom?	

The airplane is blue, and the train is yellow. Is /kas/ above that which is blue?

In Figure 3, the target word /kas/ is presented in its orthographic form above the train. The minimal pair word /kasʲ/ is presented in its orthographic form above the airplane. The answer to the question "Is /kas/ above that which is blue?" is No, since the auditory prompt tells the listener that the airplane is blue and the train is yellow and /kas/ is above the train. Each target word was presented twice, once where the answer to the question in

the trial is Yes and once where the answer to the question in the trial is No. There were 28 trials in this test.

Task D. Lexical-Access and Context

This task simulates real communication because both lexical access and the use of meaningful linguistic context are required. At the beginning of the task, the subject is instructed to answer Yes or No to the question which they will hear. In this task, the subject sees black and white line drawings of an airplane and a train. Instead of the two written words being displayed, two pictures of nonobjects will be displayed. One picture is displayed above the airplane and the other picture is above the train. A prompt will inform the subject of the color of the airplane and the train. Then the target word is embedded in a question which asks whether the target word is above the object that is a certain color. The subject must remember the colors of the airplane and the train, retrieve lexical information of the word corresponding to the pictures, and discriminate between the minimal pair in order to correctly answer the question. The color and order of the airplane and train as well as the order of the pictures are randomized in each trial. By presenting the pictures which correspond to the word, it is assumed that the subject is required to retrieve the lexical information from the lexicon. The subject is also required to understand the linguistic context to make the correct choice. Example (5) provides the auditory prompt the subject will hear while seeing a picture such as Figure 4 on the screen. In Figure 4, the target word /boz/ is presented in its picture form above the airplane. The minimal pair word /boz^{j}/ is presented in its picture form above the train. The answer to the question "Is /boz/ above that which is blue?" is Yes, since the auditory

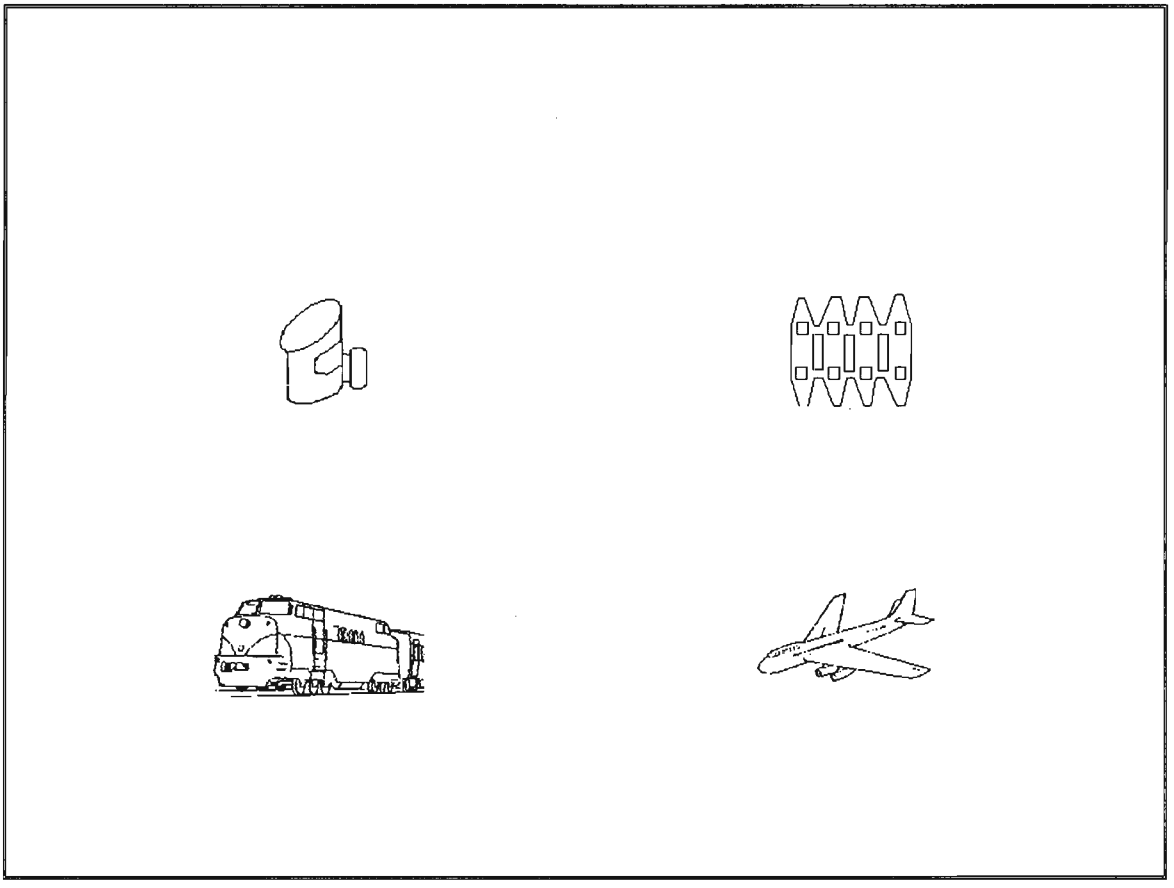


Figure 4. Screenshot of Task D; Target: /boz/

(5)	Синій	літак	та	жовтий	поїзд.	Чи
	blue-nom	airplane-nom	and	yellow-nom	train-nom.	Is
	боз	над тому,		що	синій?	
	/boz/-nom	above that-loc,		which	blue-nom?	

The airplane is blue, and the train is yellow. Is /boz/ above that which is blue? prompt tells the listener that the airplane is blue and the train is yellow and /boz/ is above the train. Each target word was presented twice, once where the answer to the question in the trial is Yes and once where the answer to the question in the trial is No. There were 28 trials in this test.

RESULTS

Responses for all the tasks were encoded as correct or incorrect. Subjects were relatively accurate in learning the novel nonwords. The mean proportion correct for the criterion test was .925 with a standard deviation of .264. As noted earlier, two subjects did not take the criterion test due to experimenter error. Of the 10 subjects analyzed, two subjects took the criterion test three times, two subjects took the criterion test twice, and six subjects took the criterion test once. Figure 5 provides the mean proportion correct for each task including the criterion test.

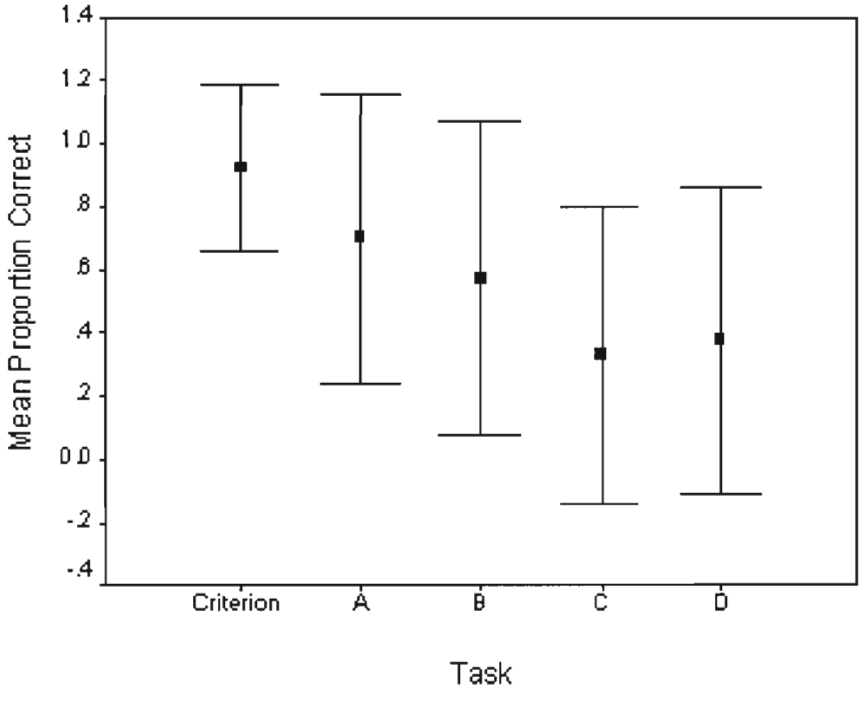


Figure 5. Mean proportion correct for each task (bars indicate +/- 1 standard deviation)

These results generally agree with the hypothesis that subjects will perceive the phonemic contrast worse on tasks that require both accessing the lexicon and processing the meaningful linguistic context. For example, subjects perceived the phonemic contrast better in Task A (No-Lexical-Access and No-Context) than in the other tasks. Analysis of variance was run with proportion correct as the dependent variable and with palatalization and task as within-subjects variables. Palatalization is defined as whether or not the consonant in the target words was palatalized. There was a main effect of task ($F(3,27) = 27.557$, $p<.005$; partial eta squared = .754), but no main effect of palatalization ($F(1,9) = .223$, $p=.648$; partial eta squared = .024) or interaction of the two ($F(3,27) = 1.463$, $p=.247$; partial eta squared = .140).

While task order was an additional between-subjects variable, only nine different orders of the four tasks were used. In addition, only one task order condition contained more than one subject. For this reason, task order was not included in the ANOVA; however, for the results of this experiment to be generalizable, more subjects should be run and the effects of task order should be investigated via ANOVA. In addition, target consonant (d, l, n, s, t, ʦ, z, and their palatalized counterparts) was not included in the ANOVA as a within-subjects variable because of the small number of observations collected from each subject.[4] Future research should either reduce the number of target consonants or increase the number of times each consonant is featured in the tasks in order to allow for a more complete investigation of possible target consonant effects. It is

[4] Weber and Cutler (2004) showed that L1 words can compete with L2 words during speech perception tasks. It is possible that perception of the phonemic contrasts in the Ukrainian nonwords were affected by similar English words. For example, the English word *feet* is similar to /fit/ and the English word *vets* (the abbreviation for veterinarian) is similar to /vets/. The results listed in Figure 6 show that any such effect was not significant.

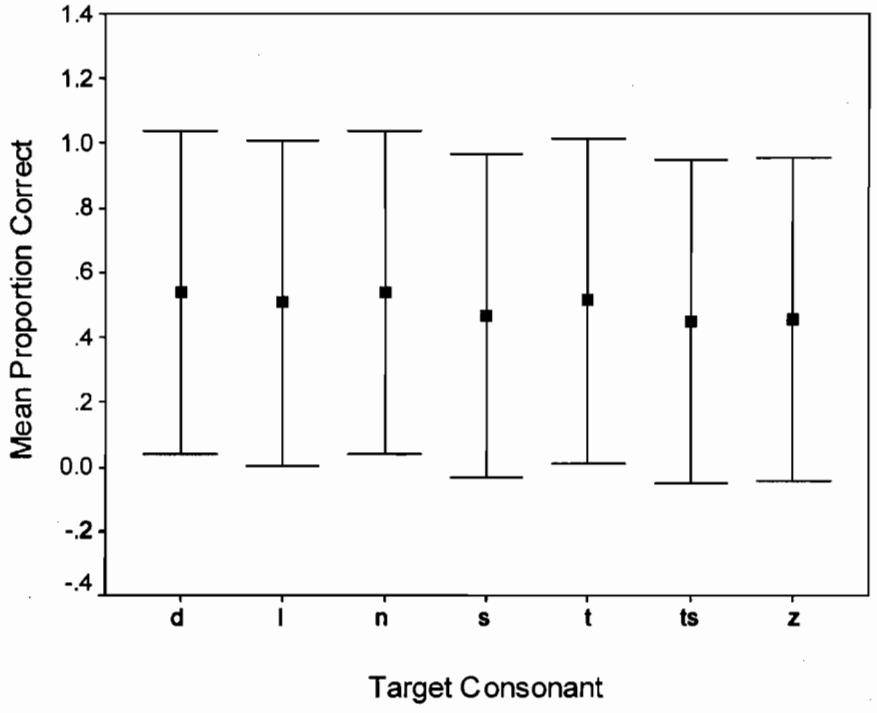


Figure 6. Mean proportion correct for each target consonant (bars indicate +/- 1 standard deviation)

possible, however, to observe the pattern of results by target consonant, as presented in Figure 6. Note that in this figure, each minimal pair is collapsed into one label (e.g. /t/ and /t^{j}/ are collapsed under the label 't').

Given the main effect of task, planned follow-up comparisons were conducted to determine the loci of the effect. In review, Task A did not require the use of lexical access or meaningful linguistic context, Task B only required the use of lexical access, Task C only required the use of meaningful linguistic context, and Task D required the use of both lexical access and meaningful linguistic context. The mean proportion correct for each task is given in Table 5. Subjects were able to perceive the phonemic contrasts significantly more often in Task A (.704) than in Task B (.571) ($F(1,9) = 9.470$, $p < .02$,

Table 5. Correct proportion for each task

Task	Mean proportion correct	Standard deviation
A	.704	.127
B	.571	.113
C	.329	.080
D	.379	.111

partial eta squared =.513), in Task C (.329) ($F(1,9) = 79.699$, $p < .005$, partial eta squared = .899), or in Task D (.379) ($F(1,9) = 27.716$, $p < .005$, partial eta squared = .755). Subjects were also able to perceive the phonemic contrasts significantly more often in Task B than in Task C ($F(1,9) = 28.582$, $p < .005$, partial eta squared = .761) or in Task D ($F(1,9) = 16.78$, $p < .004$, partial eta squared = .651). In addition, the difference between Task C and D was not significant ($F(1,9) = 1.723$, $p = .222$, partial eta squared = .161).

DISCUSSION

Table 4 shows that there is a significant difference between all the tasks except between Task C and Task D. These results justify a few conclusions. First, the need to access the lexicon reduces the ability of the subject to make use of phonemic contrasts in an auditory word recognition task. This is shown by the significant difference in the percent correct for Task A compared to Task B. Second, the requirement that learners make use of meaningful linguistic context reduces the ability of the subject to make use of the phonemic contrast. This is shown by the differences in the percent correct for Task A compared to Task C and also for Task B compared to Task D. These two conclusions hold when the percent correct for Task A is compared to Task D. Third, it seems that making use of meaningful linguistic context has a larger effect than does the requirement that the learner access the lexicon. The results show that the phonemic contrast is perceived significantly more often in Task B, which only involves lexical access, than in Task C, which only involves making use of meaningful linguistic context. This suggests that making use of meaningful linguistic context hinders the ability to perceive phonemic contrasts more than lexical access. In addition, there is no significant difference between Task C and Task D. This shows that the effect of meaningful linguistic context in perceiving the novel phonemic contrast is large enough to mask the effect of lexical access which was shown by comparing the results of Task A and Task B.

General Discussion

Curtin et al. (1998) concluded that predictable information was stored in the lexical representations created by L2 learners. However, the results of Pater (2003), which replicated the study conducted by Curtin et al. (1998), matched with other studies that supported the fact that lexical representations only include unpredictable information. Pater (2003) conducted an XAB task where subjects had to determine whether A or B matched X. There were three conditions. The first condition included three sounds (SSS). This condition did not require lexical access. The second condition introduced a sound and then two pictures (SPP). The third condition introduced a picture and then two sounds (PSS). Both the SPP and PSS conditions required lexical access. An important finding in Pater (2003) is that the results of both the lexical (PSS) and nonlexical tasks (SSS) were similar, which suggests that lexical access did not influence speech perception. On the other hand, SPP was removed from the analysis because subjects performed at chance. This shows that a lexical access condition (SPP) did increase the level of difficulty for the task. In this thesis it was assumed that the subjects had encoded the phonemic contrast in their lexical representations previous to the current experiment.

One of the results of Curtin et al. (1998) was that there is a difference between lexical and nonlexical tasks. They interpret this to mean that there seems to be a difference between what learners can perceive and what they lexically encode. This is supported by Hayes-Harb (2007), which suggests that there is an intermediate stage in the acquisition of second language phonology where learners are sensitive to a novel contrast but that they may not yet be able to represent it lexically. In addition, Hayes-Harb and

Masuda (2008) state that to correctly complete the auditory word-picture matching task, participants must have encoded information related to the contrast in their memory for the words. This shows that there are at least two levels of perception. The results of this thesis could be explained by the finding of Pater et al. (2004) that when a phonological contrast is first learned it remains partially integrated and can be lost under the processing demands of word learning. Another possible explanation could be found in the study done by Werker and Logan (1985), in which they concluded that three different levels of perception (i.e. acoustic, phonetic, and phonemic) were exhibited with three different types of tasks. The more the cognitive load increased due to an increased interstimulus interval, the less reliable the acoustic information became, and the more important the phonemic information became to the speech perception task.

Whether or not the phonemic contrast was lexically encoded was not the main issue in this thesis. It was assumed that the subjects already had a good understanding of the phonemic contrast as indicated by the proportion correct on the training test (.7911). If it is assumed that the subjects had a lexically encoded contrast due to naturalistic training, then some other factor affected the ability of the subjects to access this information in different tasks.

In this thesis then, it is possible that acoustic, phonetic, and phonemic information were available in Task A, in which lexical access and the use of meaningful linguistic context were not required, but only phonemic information in Task D, which required both lexical access and the use of meaningful linguistic context. The data in this experiment does not show what types of information were available in Task B, where only lexical access was required, and Task C, where only the processing of meaningful linguistic

context was required. However, it is possible that both Task B and Task C had less acoustic, and maybe phonetic, information available than Task A. In general, it is likely that a high-level of cognitive processing led to a phonemic level of perception, while a lower-level of cognitive processing led to a more acoustic level of perception.

An example of this is shown by the task demands of the contextual information. Ingram and Park (1998) reported that the identification and discrimination tasks in their study differed in terms of task demands on listeners. In addition they reported that phonological learning effects were more dominant than acoustic discriminability effects. Also a phonological level of signal processing was less engaged by the discrimination task. This could explain the results of this thesis if the identification tasks (Task A and B) were answered using only acoustic information, and the minimal pair tasks (Task C and D) were answered using only phonological information. This would support the idea that contextual information has a larger effect than lexical information.

This thesis is limited in that only a small sample of learners participated. This meant that subjects participated in only 11 out of a possible 24 task order conditions, and only nine of these orders were analyzed. In order to be able to generalize these results to a wider population, it will be necessary to collect data from more subjects. However, the results do point in the predicted direction, as summarized in the Conclusion section below.

CONCLUSION

The subjects who were recruited for the experiment in this thesis had learned Ukrainian in a naturalistic setting similar to the subjects in the studies conducted by Flege and colleagues, which were described previously. For this thesis it was assumed that the subjects had already lexically encoded the palatalization contrast. The first phase of the experiment was to teach the subjects new words. This word learning phase was only to familiarize the subjects with the nonwords used in the experiment. The second phase consisted of four tasks. Task A was a simple matching task where the subjects had to decide whether the auditory and written stimuli matched. Task B was an auditory word-picture matching where the subjects had to decide whether the auditory word matched the picture which was displayed. Task C was a minimal pair discrimination task embedded in meaningful linguistic context. Task D was a minimal pair discrimination task using pictures embedded in meaningful linguistic context. The subjects needed to have stored the phonemic contrast lexically in order to correctly answer the questions in Task B and Task D, while the subjects needed to have effectively processed the cognitive demands of both answering the perceptual task and processing the meaningful contextual information in Task C and Task D.

The main purpose of this thesis was to determine whether specific aspects of natural communication affect L2 speech perception. The two aspects of natural communication that were studied in this thesis were lexical access and meaningful

linguistic context. This experiment followed the pattern of recent studies in using pictures and auditory stimuli to test phonemic contrasts (Curtin et al. 1998; Pater 2003; Hayes-Harb 2007). On the other hand, this thesis is unique in that it also required subjects to process meaningful linguistic context during speech perception. The results of this thesis agree with other studies that show that as task demands increase, the ability for nonnative speakers to perceive phonemic contrasts decreases. The conclusion of this thesis is that both lexical access and meaningful linguistic context affect L2 speech perception. To study natural speech, future studies must take these factors into account.

REFERENCES

Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: the case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics, 32(2),* 233-250.

Bent, T. & Bradlow, A. R. (2003). The interlanguage speech intelligibility benefit. *Journal of the Acoustical Society of America, 114(3),* 1600-1610.

Best, C. (1994). The emergence of native-language phonological influences in infants: a perceptual assimilation model. *The development of speech perception: the transition from speech sounds to spoken words*, ed. by Judith C. Goodman and Howard C. Nusbaum, 167-224. Cambridge: MIT Press.

Boersma, P., Escudero, P., & Hayes, R. (2003). Learning abstract phonological from auditory phonetic categories: an integrated model for the acquisition of language-specific sound categories. *Proceedings of the 15th International Congress of Phonetic Sciences.*

Bradlow, A. R. & Pisoni, D. B. (1999). Recognition of spoken words by native and non-native listeners: talker-, listener-, and item-related factors. *Journal of the Acoustical Society of America, 106(4),* 2074-2085.

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America, 101(4),* 2299-2310.

Curtin, S., Goad, H., & Pater, J. V. (1998). Phonological transfer and levels of representation: the perceptual acquisition of Thai voice and aspiration by English and French speakers. *Second Language Research, 14(4),* 389-405.

Cutler, A. & Otake, T. (2004). Pseudo-homophony in non-native listening. Unpublished paper presented at the Acoustical Society of America, New York.

Escudero, P. & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition, 26(4),* 551-585.

Flege, J. E. (1984). The effect of linguistic experience on Arab's perception of the English /s/ vs. /z/ contrast. *Folia Linguistica, 18(1-2),* 117-138.

Flege, J. E. (1995). Second language speech learning: theory, findings, and problems. *Speech perception and linguistic experience: issues in cross-language research,* ed. by Winifred Strange, 233-272. Baltimore: York Press.

Flege, J. E., Frieda, E. M., Walley, A. C., & Randazza, L. A. (1998). Lexical factors and segmental accuracy in second language speech production. *Studies in Second Language Acquisition, 20(2),* 155-187.

Flege, J. E. & Liu, S. (2001). The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition, 23(4),* 527-552.

Flege, J. E. & MacKay, I. R. A. (2001). Perceiving vowels in a second language. *Studies in Second Language Acquisition, 26(1),* 1-34.

Flege, J. E., Munro, M. J., & Skelton, L. (1992). Production of the word-final /t/-/d/ contrast by native speakers of English, Mandarin, and Spanish. *Journal of the Acoustical Society of America, 92(1),* 128-143.

Flege, J. E., Takagi, N., & Mann, V. (1995). Japanese adults can learn to produce English /ɹ/ and /l/ accurately. *Language and Speech, 38(1),* 25-55.

Guion, S. G., Flege, J. E., Akahane-Yamada, R., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: the case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America, 107(5),* 2711-2724.

Guion, S. G., Flege, J. E., & Loftin, J. D. (2000). The effect of L1 use on pronunciation in Quichua-Spanish bilinguals. *Journal of Phonetics, 28(1),* 27-42.

Harnsberger, J. D., Wright, R., & Pisoni, D. B. (2008). A new method for eliciting three speaking styles in the laboratory. *Speech Communication, 50(4),* 323-336.

Hayes-Harb, R. (2007). Lexical and statistical evidence in the acquisition of second language phonemes. *Second Language Research, 23(1),* 61-90.

Hayes-Harb, R. & Masuda, K. (2008). Development of the ability to lexically encode novel second language phonemic contrasts. *Second Language Research, 24(1),* 5-33.

Heeren, Willemijn. (2004). The perceptual development of a British-English phoneme contrast in Dutch adults. *Linguistics in the Netherlands, 21(1),* 91-101.

Ingram, J. C. & Park, S. (1998). Language, context, and speaker effects in the identification and discrimination of English /r/ and /l/ by Japanese and Korean listeners. *Journal of the Acoustical Society of America, 103(2),* 1161-1174.

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(2),* 874-886.

MacKain, K. S., Best, C. T., & Strange, W. (1981). Categorical Perception of English /r/ and /l/ by Japanese Bilinguals. *Applied Psycholinguistics, 2(4),* 369-90.

Padgett, J. (2003). The emergence of contrastive palatalization in Russian. *Optimality Theory and language change,* ed. by Eric Holt, 307-335. Dordrecht: Kluwer Academic Press.

Pater, J. (2003). The perceptual acquisition of Thai phonology by English speakers: task and stimulus effects. *Second Language Research, 19(3),* 209-223.

Pater, J., Stager, C., & Werker, J. (2004). The perceptual acquisition of phonological contrasts. *Language, 80(3),* 384-402.

Piske, T., MacKay, I. R. A., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics, 29(2),* 191-215.

Pisoni, D. B. & Lazarus, J. H. (1974). Categorical and noncategorical modes of speech perception along the voicing continuum. *Journal of the Acoustical Society of America, 55(2),* 328-330.

Rubach, J. (2002). An overview of palatalization-i*. *Studies in Phonetics, Phonology and Morphology,* 8(2), 169-186.

Rubach, J. (2007). Feature geometry from the perspective of Polish, Russian, and Ukrainian. *Linguistic Inquiry, 38(1),* 85–138.

Shevelov, G. Y. (1993). Ukrainian. *The slavonic languages,* ed. by Bernard Comrie and Greville G. Corbett, 947-998. New York: Routledge.

Stager, C. & Werker, J. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature, 388(6640),* 381-382.

Strange, W. & Dittmann, S. (1984). Effects of discrimination training on the perception of /r-l/ by Japanese adults learning English. *Perception & Psychophysics, 36(2),* 131-145.

Takagi, N. (2002). The limits of training Japanese listeners to identify English /r/ and /l/: eight case studies. *Journal of the Acoustical Society of America, 111(6),* 2887-2896.

Walley, A. C. & Flege, J. E. (1999). Effect of lexical status on children's and adults' perception of native and non-native vowels. *Journal of Phonetics, 27(3),* 307-322.

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language, 50,* 1-25.

Werker, J. F. & Logan, J. S. (1985). Cross-language evidence for three factors in speech perception. *Perception & Psychophysics, 37(1),* 35-44.

Werker, J., Fennell, C., Corcoran, K., & Stager, C. (2002). Infants' ability to learn phonetically similar words: effects of age and vocabulary size. *Infancy, 3(1),* 1-30.