Designing a Study to Measure the Effects of Computer-based Minimal Pair Identification Training on L2 Aural Word Recognition and Listening Comprehension

コンピュータを利用した最小対語の認識トレーニングが第二言語の聴解力と聴覚の語彙認識に及ぼす影響の効果測定研究の設計

Kriss A. Lange

Abstract

This article describes the research design of a study to measure the effect that minimal pair identification training has on listening comprehension and the ability to recognize spoken words. An extensive literature review and theoretical background for the study are also provided. The study hopes to determine whether English learners, who practiced identifying aurally presented minimal pairs for five weeks, will improve their general listening comprehension ability and their ability to recognize aurally presented words as compared to a control group. If a causal relationship between demonstrated between these variables, there may be pedagogical implications for more focus on bottom-up skills such as phoneme identification to enhance general listening comprehension.

Listening is a fundamental skill for English language learners. It promotes the development of other language skills (Dunkel, 1991; Rost, 2002) and is an essential skill in itself for communication. However, many learners consider listening as the most difficult skill in language learning (Hasan, 2000; Kim, 2002; Graham, 2003). Managing the speed of the input, segmenting the input correctly, effectively using working memory to quickly process the input and interpreting meaning for prosodic features such as stress and intonation are some of the reasons given for its difficulty (Vandergrift and Baker 2015).

Researchers have been exploring which factors have the greatest positive impact on listening ability. One important finding is that L2 vocabulary knowledge has emerged as a strong predictor of listening comprehension ability (Staehr, 2009; Vandergrift & Baker, 2015). In Staehr's study the correlation between L2 listening comprehension and L2 vocabulary knowledge was quite strong ($r = 0.70$) but the L2 vocabulary assessments only presented the words in written form. Vandergrift and Baker used aurally presented L2 vocabulary assessments and found a weaker but still moderately strong correlation (0.49). In fact, this was the strongest correlation found among all the variables in the study including: L1 vocabulary, L2 vocabulary, L1 listening ability, auditory discrimination ability, working memory and metacognition about listening.

These studies also support findings by Mecartty (2000) which showed L2 vocabulary to be a significant predictor, explaining about 12%, of L2 listening comprehension ability. Takashima (1998) demonstrated an even larger effect with an English spoken word recognition test accounting for 57% of English listening comprehension ability for the Japanese participants. This data suggest that improving English vocabulary recognition could be an effective means to develop English listening ability.

Although vocabulary development is effective for improving listening ability, other causal factors still need to be explored. For example, auditory discrimination, the ability to receive, differentiate and process acoustic input, could have an important effect on L2 listening ability and second language development in general. Researchers have found a significant correlation between auditory discrimination and L1 language development (Tsao, Liu, & Kuhl, 2004) showing first language development at two years old is correlated with sound discrimination ability at 6 months old. According to Vandergrift and Baker (2015), "Auditory discrimination provides the crucial bottom-up information to trigger the comprehension process." Such a fundamental skill certainly would have some impact on L2 development as well.

Vandergrift (2007), in a state-of-the-art article on L2 listening comprehension research, suggested sound discrimination ability and its effect on L2 listening ability should be further researched. Wilson, Kaneko, Lyddon, Okamoto, and Ginsberg (2011) found a moderate correlation ($r = 0.37$) between the listening comprehension section of the TOEIC test and auditory discrimination ability for Japanese learners of English. Takashima (1998) also found a strong correlation ($r = 0.790$) between listening comprehension and spoken word recognition ability. Although these results are promising, research is needed to demonstrate a causal relationship between these variables rather than just a correlational one. In the discussion section Wilson et al. (2011) states, "Since sound discrimination ability is fairly strongly correlated with L2 listening proficiency, then some emphasis on the sound segment level seems justified and indeed recommended."

A number of researchers have suggested that sound and word recognition patterns of L1 phoneme acquisition are similar to those of L2 speech learning (Best & Tyler, 2007; Bundgaard-Nielsen, Best, & Tyler, 2011a, 2011b; Walley, 2007). L2 learners first detect and analyze the prosodic patterns of the second language before recognizing vocabulary words and building their emerging lexicons. As those lexicons grow, the learner is challenged to recognize and identify more and more phonetically similar words such as minimal pairs. This leads to better phoneme discrimination ability and more advanced L2 listening comprehension ability. This view of L2 speech learning (i.e., a perception-first view) assumes that phoneme perception drives the learning process leading to production ability (Flege, 1995, 2003; Kuhl, 2000).

The Speech Learning Model (SLM), developed by Flege (1995), also provides useful insights on how new L2 speech sounds are acquired. The model claims that the learner cannot perceive the distinction between two similar phonemes because the sounds have been combined into a single category that has been influenced by the learner's L1. Therefore, phonologically similar sounds, rather than differing sounds pose more of a challenge to accurate perception due to this phenomenon of lumping distinct sounds together. For example, the English /l/ - /r/ distinction is often perceptually categorized by Japanese English learners as the nearest L1 counterpart (Japanese tap /ɾ/) for both of these phonemes thereby masking the difference (Guion, Flege, Ahahane-Yamada, & Pruitt, 2000).

It is logical to assume that accurately perceiving and discriminating English phonemes into their separate categories would help improve English spoken word recognition and overall listening comprehension. Research has shown auditory training to be effective for developing aural identification of English phonemes (e.g., Bradlow, Pisoni, Akahane-Yamada & Tohkura, 1997; McClelland, Fiez & McCandliss, 2002; Wang & Munro, 2004). Participants in these studies learned to perceive and identify phonemes more accurately through auditory training. Auditory training refers to "training the ears to hear speech features and sound distinctions that do not exist in the L1."(Grant, 2014). For instance, in the study by McClelland et al. (2002), participants heard the words "Load" or "Road" and indicated if the initial consonant was /l/ or /r/ by pressing the corresponding letter on a keyboard. The group that received no feedback made little progress. However, the group which got immediate feedback could determine if their choice was correct or not and began to identify the phonemes with better accuracy. A similar procedure was used in Wang and Munro (2004) to train participants to perceive English vowel sounds presented with minimal pairs. Again, corrective feedback was used leading to significant gains in perception accuracy. A retention test given three months after the treatment showed identification scores had dropped only slightly from the post-test values showing the changes in perception were long-lasting.

There is also evidence that auditory training alone has a positive impact on pronunciation output even without explicit pronunciation instruction (Bradlow, et al., 1997; Derwing, Munro & Wiebe, 1998). These studies showed that treatment groups not only learned to perceive phonemes with more accuracy but also learned to pronounce the phonemes better despite receiving no explicit instruction. According to Bradlow et al. (1997), "…learning in perception and production are closely linked, since perceptual learning generally transferred to improvement in production."

Auditory training has been shown to enhance perceptual learning of phonemes and pronunciation accuracy. However, there seems to be very little research investigating how aural word recognition and listening comprehension develop as a result of perceptual learning through auditory training. English has 14 vowels, 5 diphthongs and 24 consonants (Small, 2012) whereas there are only 5 vowels and 14 consonants in Japanese (Tsujimura, 1996). Japanese English learners must somehow train their ears to hear these phonemes in order to accurately comprehend spoken English. If English phonemes aren't accurately perceived, overall listening comprehension may be hindered. The purpose of this research is to investigate the effect that auditory training has on listening comprehension abilities.

Research Questions

What effect does auditory training have on the following skills: (1) minimal pair identification, (2) sound discrimination, (3) aural word recognition, and (4) listening comprehension compared with a control group that practiced audio-assisted extensive reading?

Research Methods

I will begin conducting this research project with the first-year Nursery Education majors taking my English Fundamentals II course, offered in the second semester (from October 1st, 2016) at the University of Shimane Junior College. This course was chosen because over 50 students enroll each year, providing a larger sample size than any of the other courses I teach. Also, this is the only English course those students will take during the semester so influence from other English courses will be reduced.

According to the results of the Pre-tests, I plan to divide the students into Treatment and Control groups with roughly 25 students in each group. Both groups will be given a four-part pre-test and survey administered via the learning management system Moodle. The first test will be a minimal pair identification test to establish a baseline for this ability before students in the treatment begin the five weeks of auditory training using the same minimal pairs. Next, they will take the sound discrimination test to measure their ability to perceive most of the vowel and all of the consonant segments in English. After that, they will take the Listening Vocabulary Levels Test (LVLT) to measure their ability to recognize words aurally. Participants will also take a text version of the LVLT to determine word recognition in print compared with aural recognition. The final test will be the listening section of the TOEIC IPT exam to measure overall listening comprehension. Lastly, I will ask the class to complete a survey to learn more about their experiences with learning English and self-estimations of their listening skills. Each part of the pre-test will be discussed in more detail below.

Minimal Pair Identification Test

This test will be used to assess the participants' ability to identify an aurally presented word in a minimal pair. For example, students will hear "think" and then see the question written, "Did you hear sink or think?" and they will select their answer. In order to select the phonemes to use for this study I referred to Saito (2011) which identifies eight English key segmentals, /æ, f, v, θ, ð, w, l, ɹ/ that were selected by 48 experienced native Japanese English teachers as difficult for Japanese English learners to produce. In addition to these, I have also included /s, ʃ, ɛ, ɪ/ in order to form contrasting minimal pairs and to investigate other phonemes such as the /s/ vs. /ʃ/ distinction that I have noticed to be troublesome for students. The test will use minimal pairs of words in which one of the words contains one of the 12 segmentals listed above, and the other word differs by only one phoneme in the same word position. For example, to assess the participant's ability to identify /æ/, I will use “cat” and “cut” with the target phoneme in the same middle position of the word. All of the minimal pairs can be seen in Table 1.

Table 1: 16 Segmentals and Two Sets of Minimal Pairs to be used in the Minimal Pair Identification Test.

English Segmentals	Minimal Pair 1	Minimal Pair 2
/æ/ - /ɛ/	Cat - Cut	Hat - Hut
/f/ - /v/	Safe - Save	Leaf - Leave
/θ/ - /f/	Thin - Fin	Thirst - First
/ð/ - /s/	That - Sat	They - Say
/w/ - /v/	Wine - Vine	West - Vest
/l/ - /ɹ/	Light - Right	Lock - Rock
/s/ - /ʃ/	Sea - She	Sell - Shell
/ɛ/ - /ɪ/	Dead - Did	Bear - Beer

Recognition ability of these minimal pairs will be tested in the pre and post- test and they will be used for auditory training in the treatment group during the course.

Sound Discrimination Test

The sound discrimination test will be based on the test described in Wilson et al. (2011) and will be administered using the multiple-choice quiz format in the open-source course management system Moodle. Test instructions will be provided in Japanese and students will have sufficient time to read them and ask questions before starting. Students will be asked to listen to each item only once. Each item will be a nonsense syllable – either CV or VC, with the target sound naturally occurring in English. For consistency, if a consonant (C) was the focus; the vowel used will always be /a/. If a vowel (V) was the focus, the consonant used will always be /p/. The nonsense syllables can be seen in Table 2. All tokens will be pre-recorded by the author in a quiet environment. After listening to an item, the subject will choose one of four answers – the one that contained the sound s/he heard. Both question order and the order of the four answer choices will be randomized for every subject.

The answers will be given in the form of very common English words, with one sound underlined (see Table 3). For example, in the case of the syllable /pa/ where students are instructed (in Japanese) to choose the word that contains the same consonant sound, the answer choices are “pin”, “been”, “fit”, and “voice”. Distracter answer choices are chosen to have the most frequent perceptual confusions as target sounds.

Table 2: 82 nonsense syllables for the sound discrimination test.

Type of syllable	Nonsense syllables used
Pre- /a/ consonant (22)	*/pa, ba, ta, da, ka, ga, ma, na, va, fa, θa, ða, sa, za, ʃa, ha, tʃa, dʒa, wa, ja, ra, la/*
Post- /a/ consonant (21)	*/ap, ab, at, ad, ak, ag, am, an, av, aŋ, af, aθ, að, as, az, aʃ, aʒ, atʃ, adʒ, ar, al/*
Pre- /p/ vowel (10) + (3) diphthongs	*/ip, ep, ɪp, ɛp, æp, up, ʊp, op, ʌp, ap, aɪp, aʊp, ɔɪp/*
Post- /p/ vowel (10) + (3) diphthongs	*/pi, pe, pɪ, pɛ, pæ, pu, pʊ, po, pʌ, pa, paɪ, paʊ, pɔɪ/*

Table 3: All words that will appear as answer choices in the sound discrimination test.

Question type	Answer choice words used
Pre- /a/ consonant	pin, been, toss, done, kiss, guess, miss, nice, voice, fit, thin, those, sit, zoo, shop, hit, check, juice, wet, yes, rice, like
Post- /a/ consonant	zip, web, sit, need, sick, log, him, seen, love, sing, wife, tooth, father, nice, cause, wish, pleasure, peach, judge, four, kill
Pre- /p/ vowel	meet, take, kick, set, cat, food, took, coat, luck, hot, nice, house, coin
Post- /p/ vowel	meet, take, kick, set, cat, food, took, coat, luck, hot, nice, house, coin

Listening Comprehension Test

In this study I define L2 listening comprehension according to Buck (2001):

> (it is) "the ability to 1) process extended samples of realistic spoken language, automatically and in real time; 2) understand the linguistic information that is unequivocally included in the text; and, 3) make whatever inferences are unambiguously implicated by the content of the passage."

In order to measure the participants' listening comprehension, I plan to use an abridged version of the listening section of the TOEIC IP exam. The TOEIC IP Test (TOEIC = Test of English for International Communication; IP = Institutional Program) consists of 100 Listening questions and 100 Reading questions and lasts for 2 hours. The maximum total score for the listening test is 495 points. Pan (2010) summarized the TOEIC listening section as follows: "The listening tasks consist of four parts: (1) choosing the best description that matches the photograph, (2) responding to one short question or statement, (3) choosing the best response to the question from a conversation, and (4) choosing the best response to the question from a short talk." This test is well-known to Japanese students and provides a reliable measure of overall English listening comprehension.

Aural Word Recognition Test

The LVLT, created by McLean, Kramer and Beglar (2015), tests the aural lexical knowledge of the 5000 most frequent words by 1000-word frequency bands. It is based on the word frequency list by Nation (1990) of the most frequently occurring words in the BNC/COCA corpus along with words from the Academic Word List by Coxhead (2000). McLean *et al*. (2015) designed the LVLT to be completed in about 30 minutes. The test, audio files and the text version of the LVLT which is called the New Vocabulary Levels Test (NVLT), have all been made freely available online for teaching and research. Within three weeks of taking the LVLT, I plan to give all the participants the NVLT to determine which words each student could recognize in text but not aurally. I will focus on these words to determine whether auditory training influences aural word recognition of previously unrecognized words on the post-test.

The survey will explore student perceptions of their English levels, English learning experience, their perceived importance of learning English and a self-assessment of their level of accuracy for recognizing English phonemes and words aurally.

Procedures for Treatment and Control Groups

The Control group will spend 30 minutes during class doing audio-assisted extensive reading for 5 weeks. Students will be able to access the audio and text through the Moodle course management website and listen while reading with headphones during class. This task was selected for the control group to investigate whether students would make gains on the post-test implicitly through listening and reading. Also, audio-assisted extensive reading has been shown to improve reading rates and comprehension levels (Chang & Millet, 2015) thus allaying ethical concerns of improper use of class time for the control group.

The Treatment group will undergo auditory training by identifying the aurally presented word in minimal pairs via Moodle during the same 30 minutes that the control group is reading and listening. For example, students will hear a word such as “she” and see the minimal pair “sea” and “she” displayed on the screen with its Japanese definition. Then they will select the word they heard and get feedback as to which word was actually spoken. The in-class activities for the control and experimental groups will only take place during a 30-minute period of class time to ensure equal time-on-task for both groups for five weeks.

Following the treatment period, the post-tests and post-survey, which are identical to the pre-tests, will be given to all of the students and the results will be analyzed. I speculate that after auditory training, the treatment group’s scores for (1) minimal pair identification, (2) sound discrimination, (3) aural word recognition, and (4) listening comprehension will exceed those of the control group. In addition, the number of words students in the treatment group could recognize in text on the NVLT, but could not recognize when presented aurally on the LVLT will decrease. I also expect to find a correlation between the listening comprehension test scores and results of the aural word recognition test. The proposed reason for increases for the treatment group is that through auditory training, students were able to form distinctions between English phonemes which enabled them to construct a more accurate mental phonemic representation of English sounds which thereby improved aural word recognition and listening comprehension.

References

Best, C., & Tyler, M. (2007). Nonnative and second-language speech perception. In O. Bohn & M. Munro (Eds.), *Language experience in second language speech learning: In honour of James Emil Flege* (pp. 13–34). Amsterdam: John Benjamins.

Bradlow, A., Pisoni, D., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/. *Journal of the Acoustical Society of America, 101*, 2299-2310.

Bundgaard-Nielsen, R., Best, C., & Tyler, M. (2011a). Vocabulary size is associated with second-language vowel perception performance in adult learners. *Studies in Second Language Acquisition*, 33, 433–461.

Buck, G. (2001). *Assessing listening*. New York: Cambridge University Press.

Bundgaard-Nielsen, R., Best, C., & Tyler, M. (2011b). Vocabulary size matters: The assimilation of second-language Australian English vowels to first-language Japanese vowel categories. *Applied Psycholinguistics*, 32, 51–67.

Chang, A. C.-S., & Millet, S. (2015). Improving reading rates and comprehension through audio-assisted extensive reading for beginner learners. *System*, 52, 91–102.

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34, 213–238.

Derwing, T., and Munro, M., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, 48(3), 393–410.

Dunkel, P. (1991). Listening in the native and second/foreign language: Toward an integration of research and practice. *TESOL Quarterly*, 25(3), 431–457.

Flege, J. (1995). Second language speech learning: Theory, findings and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). Timonium, MD: York Press.

Flege, J. (2003). Assessing constraints on second-language segmental production and perception. In A. Meyer & N. Schiller (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 319–355).

Graham, S. (2003). Learner strategies and advanced level listening comprehension. *Language Learning Journal*, 28, 64– 69.

Grant, L. (2014). *Pronunciation myths*. Ann Arbor: University of Michigan Press.

Guion, S., Flege, J., Ahahane-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, 2711–2725.

Hasan, A. (2000). Learners' perceptions of listening comprehension problems. *Language, Culture and Curriculum*, 13, 137–153.

Kim, J. (2002). Affective reactions to foreign language listening retrospective interviews with Korean EFL students. *Language Research*, 38, 117–151.

Kuhl, P. K. (2000). A new view of language acquisition. Proceedings of the National Academy of Science, 97, 11850–11857.

McClelland, J., Fiez, J., & McCandliss, B. (2002). Teaching the /r/ - /l/ discrimination to Japanese adults: Behavioral and neural aspects. *Physiology and Behavior*, 77, 657-662.

McLean, S., Kramer, B., & Beglar, D. (2015). The creation and validation of a listening vocabulary levels test. *Language Teaching Research*, 19(6), 741-760.

Mecartty, F. (2000). Lexical and grammatical knowledge in reading and listening comprehension by foreign language learners of Spanish. *Applied Language Learning*, 11, 323–348.

Nation, P. (1990). *Teaching and learning vocabulary*. New York: Newbury House Harper Row.

Pan, Y.-C. (2010). Enhancing students' communicative competency and test-taking skills through TOEIC preparatory materials. *TESOL Journal*, 3, 81-91.

Rost, M. (2002). *Teaching and researching listening*. Harlow: Longman.

Saito, K. (2011). Identifying problematic segmental features to acquire comprehensible pronunciation in EFL settings: The case of Japanese learners of English. *RELC Journal*, 42(3), 363–378.

Small, L. H. (2012). *Fundamentals of Phonetics: A practical guide for students*. (3rd ed.). Upper Saddle River, NJ: Pearson Education.

Staehr, L. S. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition*, 31, 577–607.

Takashima, H. (1998). Accuracy of spoken word recognition as a predictor of listening comprehension for Japanese learners of English. *Annual Review of English Language Education in Japan*, 9, 87–95. http://ci.nii.ac.jp/naid/110008512160

Tsao, F.M., Liu, H. M., & Kuhl, P. K. (2004). Speech perception in infancy predicts language development in the second year of life: A longitudinal study. *Child Development*, 75, 1067–1084.

Tsujimura, N. (1996). *An introduction to Japanese linguistics*. Oxford: Blackwell.

Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching*, 40(3), 191–210. http://doi.org/10.1017/S0261444807004338

Vandergrift, L., & Baker, S. (2015). Learner variables in second language listening comprehension: An exploratory path analysis. *Language Learning*, 65(2), 390–416. http://doi.org/10.1111/lang.12105

Walley, A. (2007). Speech learning, lexical reorganization, and the development of word recognition by native and non-native English speakers. In O. Bohn & M. Munro (Eds.), *Language experience in second language speech learning: In honour of James Emil Flege* (pp. 315–330). Amsterdam: John Benjamins.

Wang, X., & Munro, M. (2004). Computer-based training for learning English vowel contrasts. *System*, 32, 539-552.

Wilson, I., Kaneko, E., Lyddon, P., Okamoto, K., & Ginsburg, J. (2011). Nonsense-syllable sound discrimination ability correlates with second language (L2) proficiency. In XVII International Congress on Phonetic