

Korean listeners' perception of L2 English phoneme contrast*

Gwanhi Yun
(Daegu University)

Yun, Gwanhi. 2014. Korean listeners' perception of L2 English phoneme contrast. *Studies in Phonetics, Phonology and Morphology* 20.2. 161-185. The current study investigates how Korean L2 listeners identify English minimal pairs contrasting in voicing and manner and attempts to testify the essential tenets of the Perceptual Assimilation Model. Another purpose of our research is to explore other factors which might affect the identification patterns within the identical contrast categories proposed by PAM. Identification tasks were performed with target vowel and consonant contrasts within real word and nonce word conditions. Identification accuracy and reaction time were measured to evaluate the relative difficulty of L2 sound identification. First, we found that Korean listeners encountered the most difficulty in identifying L2 English minimal pairs in the Single Category contrast, followed by those in the Category Goodness contrast, and those in the TC contrast were most accurately and rapidly identified in accordance with the predictions of the PAM model. Next, it was shown that there were varying levels of difficulty in identification tasks within the TC and SC contrast. This finding suggests that PAM is not sufficient to predict L2 identification accuracy and other factors should be taken into consideration, including the Speech Learning Model, and other phonological factors such as manner of articulation, the vocalicity of the segments, individual contrast type, etc. (Daegu University)

Keywords: cross-language speech perception, similarities between L1 and L2 sounds, perceived (dis)similarities, Perceptual Assimilation Model.

1. Introduction

For the decades there have been numerous studies on how L2 listeners perceive the cross-language phonetic contrast in L2 sounds or between L1 and L2 phonemes and on how they produce or acquire L2 contrast with respect to L2 development (Best 1993, 1994, 1995, Flege 1987, 1995a,b). Two of the most influential models pertaining to these issues have been put forward and developed: Perceptual Assimilation Model (henceforth PAM; Best 1993, 1994, 1995) and Speech Learning Model (henceforth SLM; Flege 1987, 1995a, b). While the traditional contrastive analysis model argues that the difference between L1 and L2 grammatical items makes it difficult to acquire L2, PAM and SLM suggest that the perceived phonetic (dis)similarities between L1 and L2 speech or in L2 contrast determine the feasibility of L2 acquisition. To be more specific, according to SLM and PAM, the greater perceived distance from L1 to L2 sounds along with more

* This research was supported by the Daegu University Research Grant, 2011. I would like to thank three anonymous reviewers for their useful comments and feedback. All errors remain my own.

exposure to L2, the more likely L2 learners are to forge a new phonetic category. On the contrary, the smaller phonetic difference between L1 and L2, the more difficulty L2 learners will likely have in acquiring L2 sounds. Flege's (1987, 1995, 2002) SLM also proposed two influential factors, i.e., age of learning (AOL), and perceived phonetic distance with respect to the development of a new L2 phonetic category. The SLM assumes that since L1 and L2 sounds exist in the same and one phonological space and the mechanism of L1 acquisition remains intact over the life span, L1 and L2 sounds constantly interact with each other, giving rise to phonetic assimilation or dissimilation. Phonetic assimilation refers to the process whereby L2 sounds are identified with its close L1 sound because of the close perceived phonetic distance. Conversely, phonetic dissimilation occurs when a new category for L2 is formed because of the far perceived distance between L1 and L2 sounds. In this vein, the role of perception in L2 production has also been spotlighted (Rochet 1995, Flege et al. 1997). For example, Rochet (1995) suggests that imperfect perception of a new sound contrast might induce deficient production of the relevant sounds.

However, these two models have constantly faced criticism on what the metric for acoustic or perceived similarity is. It is common agreement that it is difficult to quantify the phonetic similarity.¹ Furthermore, Strange et al. (2004) also claimed that unlike SLM, acoustic similarity could not always result in perceptual similarity. However, Gilichinskaya and Strange (2010) showed that acoustic similarity might be a good predictor of the phonetic distance between L1 and L2 sounds.

Since our study explores Korean L2 listeners' perception of English consonants and vowels, we will briefly review L2 listeners' perception of vowel and consonant contrast, respectively.

1.1 L2 listeners' perception of vowel contrast

As SLM predicts, less advanced L2 listeners who have less experiences in L2 tend to resort to one specific acoustic feature over another. For instance, Flege and Bohn (1989) showed that native Spanish listeners identified English vowels such as /i/ and /ɪ/ in *beat* and *bit* on the basis of vowel duration rather than spectral qualities such as F1 values. They suggested that the former is more readily accessible to L2 listeners than the latter and thus is likely to be memorized in their mental representations.

Another reason for difficulty for L2 perception is that two L2 vowels are identified or classified as one L1 vowel category (Flege 1995b, Best 1995).

¹ In SLM, phonetic acoustic similarity is adopted, whereas in PAM, phonetic distance is gauged by "the spatial proximity of constriction locations and active articulators and by similarities in constriction degree and gestural phasing" (Best 1995: 194). Adopting "direct realist perception theory", PAM posits that L2 speakers tend to identify L2 sound by way of gestural similarity to its close L1 sound. Furthermore, James (1984) proposes three metrics for phonetic similarity such as gesture, acoustics features, and abstract phonological ones.

For example, Flege (1995b) interpreted Major's (1987) finding such that native Spanish listeners might have identified both L2 English vowels /æ/ and /ɑ/ as Spanish /a/ because of a limited Spanish vowel inventory (/i e a o u/). Likewise, Flege (1995a) found that Korean speakers showed great confusion, failing to discriminate English vowel contrasts /i/-/ɪ/, /æ/-/ɛ/, /u/-/ʊ/, /ɑ/-/ʌ/, and /ɛ/-/ɪ/. To be specific, English /i/ was wrongly identified as /ɪ/ by 33%, and the reverse was by 23%, indicating that the English vowels can be perceived equally well as Korean /i/. Such a case might represent Single Category pattern whereby two L2 categories can be perceptually assimilated to one L1 category with a similar degree of perceptual distance according to Best's (1995) PAM. It is highly plausible that L1 Korean does not exhibit those two categorical contrasts and thus English /i/-/ɪ/ can be classified as Korean /i/, and English /æ/ and /ɛ/ cannot easily be discriminated because the corresponding Korean vowels /æ/ and /ɛ/ have undergone merge in contemporary Korean.

An interesting example of the claim that acoustic similarity between L1 and L2 sound can influence L2 sound categorization comes from Escudero et al.'s (2012) perception study. The authors investigated whether the different dialects of L1 Dutch affect the perceptual patterns for L2 English vowel contrast /ɛ/-/æ/. First, they showed that speakers of two dialects (the North Holland, Flemish) produced their vowels differently with respect to F1 and F2 values. Second, they found that L2 listeners of different Dutch dialects yielded different perceptual categorization. To be specific, English /ɛ/ was 100% classified as Dutch /ɛ/ by the North Holland dialect, whereas the Flemish listeners identified with /ɛ/ by 70%. For English /æ/, the former dialect listeners classified it as either Dutch /ɑ/ (50%) or /ɛ/ (50%) while it was identified as Dutch /ɛ/ by 100% by the latter listeners. This finding provides further support for the claim that the perceptual classification of L2 sounds are contingent upon acoustic similarity between the sound of one dialect and L2 sound.

A bulk of previous studies cited above lead us to suggest that the extent and manner of L2 sound perception vary as a function of multiple factors: (i) age of L2 acquisition, (ii) the amount of L2 exposure, (iii) perceived acoustic similarity between L1 and L2 sounds, (iv) similarity of articulatory gestures and/or acoustic parameters between L1 and L2 sounds, etc. Given this background, the present study investigates how Korean L2 learners of English perceive tense-lax English vowel contrasts.

1.2 L2 listeners' perception of consonant contrast

Another goal of the present study is to examine how Korean learners of English perceive consonantal contrasts in voicing (voiced-voiceless), manner (stop-fricative), place (/s/-/θ/) and liquids (/l/-/r/) in L2 English, and thus previous studies will be briefly overviewed concerning the relevant factors in L2 consonant categorization. Unlike the traditional contrastive analysis, the

aforementioned SLM posits an interesting claim that the farther the perceived distance between L1 and L2 sounds, the more likely the L2 sound is to be established as a new category. Many pieces of supportive evidence have been unfolded regarding Japanese L2 learners' acquisition of English /l-/r/ contrast. For example, Sekiyama and Tohkura (1993) and Takagi (1993) showed that Japanese learners produce English /r/ more accurately rather than /l/ especially in syllable-final position. This is offered an explanation that English /l/ is perceptually more similar to Japanese /r/ than English /ɹ/, making it comparatively easier to establish English /ɹ/. This might lead us to speculate that perceived phonetic similarity might be indexed in a variety of levels such as allophones, phonemes, or phonotactic distribution (Strange and Shafer 2008).

Another example of difficulty of L2 consonantal category formation is seen in the case of mastering English voiceless stops (Flege and Eefting 1987). Flege and his colleagues found that L2 adult learners whose L1s are Spanish, Dutch, and Italian produce English voiceless stops /p, t, k/ with shorter VOTs than native English speakers. This finding offers supportive evidence for SLM's equivalence classification. Specifically, it is posited that since some L2 consonant is incorrectly perceived and classified into a similar L1 category, blocking an accurate L2 category formation and leading to incorrect production (Flege and Munro 1994). Inappropriate mapping between L1 and L2 categories due to seemingly perceived similarity can be a likely factor in the formation of L2 category. Such an incorrect linkage between L1 and L2 consonants might plausibly account for the impressionistic observation that Korean L2 learners often produce English word-initial /ð/ as /d/ or Korean lax /t/ and English /θ/ as /s/ or Korean tense /s'/. That is why the current study attempts to reveal the identification patterns for L2 stop-fricative contrasts (e.g., /b-v/, /p-f/, /d-ð/).

As is known, unlike SLM aiming to explain mapping L1 to L2 category, Best and her colleagues have focused on perception of L2 category contrasts. Under PAM, based on the gestural similarity between L1 and L2, two L2 sounds can be assimilated to various categories in L1 or as nonspeech sounds: (i) Two-Category Assimilation (TC Type) where two L2 sounds are assimilated to two different L1 categories (e.g., English /t/ => Korean /t^h/, English /d/ => Korean /t/), (ii) Category-Goodness Difference (CG Type) where two L2 sounds are assimilated to two L1 categories with different distances, and (iii) Single-Category Assimilation (SC Type) where two L2 sounds are assimilated to a single L1 category with equal perceptual distance. Many of Best and her colleagues' studies have shown that L2 sound contrasts are perceived according to the perceptual mapping types they proposed. For instance, it was found that Zulu clicks were perceived as nonspeech sounds by English speakers (NA type) because the articulatory gestures of double-stop and suction release for clicks did not exist for English consonants. Furthermore, voiceless and voiced lateral fricatives in Zulu were categorized into two English consonants (TC type), and a velar voiceless aspirated and

ejective stops were identified as English /k/ with different degrees of goodness (CG type) (Best et al. 1988).

More recently, Pilus (2005) conducted a perception experiment to examine how Malay speakers of English perceive English word-final obstruent voicing contrasts and to test PAM's category mapping hypotheses. As expected, English /t/-/d/ contrast was most accurately identified by 92.5% accuracy because Malay also has /t/ and /d/ in phoneme inventory (TC type), and another TC type, /f/-/v/ contrast was also quite well perceived (85.4%). However, /s/-/z/ voicing contrast was least accurately identified (79%), which is attributed to its CG type where Malay has /s/ but no /z/. This finding was interpreted to support the PAM's claim that a TC contrast is easier to categorize for L2 sounds than the other types.

Park and de Jong (2008) conducted a comprehensive labeling perception task to look into what Korean L1 categories English consonants are perceived as. For instance, English /p/ was mapped onto Korean aspirated /p^h/ by 95%, whereas English /b/ was perceptually mapped onto Korean tense /p'/ by 46%, Korean lax /p/ by 41%, and Korean aspirated /p^h/ by 85%, etc. It was shown that it is quite possible to predict the patterns of mapping L2 stops onto L1s only by referring to L1 categories. Their focus, however, has been limited to the category mapping of English-Korean consonants.

In sum, previous studies attempted to look into the mapping of L2 categories of a single segment (SLM) or relations between two categories (PAM) by employing many L2 and L1 segments.

1.3 Purpose of this study

Despite numerous studies on the category mappings between L1 and L2 sounds, much attention has not been paid to a simpler or more comprehensive look into the identification probability with regard to L1 and L2 sounds. Thus one goal of the current research is to investigate Korean speakers' perceptual mapping patterns of English L2 segments by examining the L2 identification accuracy and reaction time when it took respond to identify L2 segments.

Let us examine the rationale behind the assimilation patterns for L2 English sound pairs categorized for the current study. The categorization is based on previous identification or labeling studies conducted by Jun (2003), Park and de Jong (2008: 709), de Jong and Cho (2012: 364-366) and Son (2008). First, the English sound pairs under the TC below were labeled as two distinct Korean phonemes with the great degree of identifiability. For instance, according to Park and de Jong (2008), English /p/ was labeled as Korean /p^h/ by 95% and English /b/ as Korean /p'/ by 46% or as /p/ by 41%. English /t/ was identified as Korean /t^h/ by 98% and English /d/ as Korean /t/ by 84% with no overlap of labeling. As for the perception of English /k-g/, this pair was categorized as TC based on the expectation that /k/ would be likely to be identified mostly as Korean /k^h/, and /g/ as Korean /k/ compared

with the cases of English /p-b/ and /t-d/. Next, according to de Jong and Cho (2012), English /s/ in word initial position was perceived as Korean /s'/ by 88% and English /z/ as Korean /tʃ/ by 95%, which makes it sense that this contrast is classified as TC. Additionally, English /f/ was adapted as Korean /p^h/ by approximately 54% and as /p'/ by 22%, and English /v/ as Korean /p/ by 65%. As for /tʃ-dʒ/ pair, English /tʃ/ is expected to be perceived as Korean /tʃ^h/, and English /dʒ/ as Korean /tʃ/.

Next, let us examine the pairs corresponding to CG type. First, English /s/ in word initial position was perceived mostly as Korean /s'/ by 88% and English /θ/ as Korean /s'/ by 40% (de Jong and Cho 2012: 364). Second, English /d/ was adapted as Korean /t/ by 84%, and English /ð/ as Korean /t/ by 78%. Next, English /p/ and /f/ were both identified as Korean /p^h/ by different accuracy, respectively 93%, and 54%. Finally, English /b/ and /v/ in word initial position were perceived substantially as Korean /p/, respectively by 41% and 65%. Put differently, all of these English sound pairs mentioned here could be perceptually identified as a single L1 category with substantially different proportion, and thus can be classified as CG type.

Finally, the classification of three vowel pairs into SC type was done on the basis of phonological patterns of English and Korean. Since L2 English phonology has tense-lax high vowels (/i-ɪ/, /u-ʊ/) whereas Korean has only single corresponding high vowels (/i/ and /u/), it has been known that Korean learners face serious challenges in identifying English vowel pairs (Son 2008).² Son's (2008) study showed that English /i/ was perceived as /i/ by 40% and as /ɪ/ by 43% especially by inexperienced Korean learners. Taking together her results and phonological systems of Korean and English, these English pairs can be assimilated to Korean /i/ with possibly similar degrees of difficulty. Thus, these pairs can be categorized into SC type. As for English /e-æ/ pair, although Korean has distinct phonemes like /e-æ/, the pair has almost merged into one in modern Korean at phonetic level, and thus, English contrast /e-æ/ is likely to be assimilated either to Korean /e/ or to /æ/. Son's (2008: 59) study shows that English /e/ is identified as /e/ by 41% and as /æ/ by 34%, indicating that inexperienced Korean speakers seem to be confused with identification of this contrast. Furthermore, English /ʊ/ was identified as /u/ only by 36% and as /u/ by 32%. Hence, this result leads to plausible rationale that all of the three vowel contrasts in English can be perceptually assimilated into a single Korean vowel category with almost equal perceptual difficulty, lending support to the categorization into SC type.

² As pointed out by an anonymous reviewer, a limitation of the current study is that the identification accuracies cited from previous studies, in particular, Park and de Jong (2008), and de Jong and Cho (2012) refer to the percentages for word-initial position rather than word-final position the current study focuses on.

(1) Category types of the experimental stimuli

| Contrast | Category type | Eng-Kor | Contrast | Category type | Eng-Kor |
|-----------|---------------|--------------------------------------|----------|---------------|-----------------------------------------------------------|
| /p/-/b/ | TC | /p/→/p ^h / /b/→/p/ | /s/-/θ/ | CG | /s/→/s'/(88%) /θ/→/s'/(40%) |
| /t/-/d/ | TC | /t/→/t ^h / /d/→/t/ | /d/-/ð/ | CG | /d/→/t/ (84%) /ð/→/t/ (78%) |
| /k/-/g/ | TC | /k/→/k ^h / /g/→/k/ | /p/-/f/ | CG | /p/→/p ^h /(95%) /f/→/p ^h / (54%) |
| /s/-/z/ | TC | /s/→/s'/ /z/→/c/ | /b/-/v/ | CG | /b/→/p/ (41%) /v/→/p/ (65%) |
| /f/-/v/ | TC | /f/→/p ^h / /v/→/p/ | /i/-/ɪ/ | SC | /i, ɪ/→/i/ |
| /tʃ/-/dʒ/ | TC | /tʃ/→/tʃ ^h / /dʒ/→/tʃ/ | /ɛ/-/æ/ | SC | /ɛ, æ/→/ɛ/ or /æ/ |
| /l/-/r/ | CG | /l, r/→/l/ ³ | /u/-/ʊ/ | SC | /u, ʊ/→/u/ |

Another purpose of this research is to find out other potential phonological factors other than PAM-based category types which might affect the degree of difficulty of identification such as voicing of consonants, manner of articulation, or vowel type. This probe might reveal a more comprehensive picture of L2 category identification or complement any L2 sound perception models. It is based on speculation that there might be other factors to determine the difficulty of identification of L2 contrast even in the identical category type, i.e., Two Categories, Category Goodness, or Single Category as suggested by Pilus (2005).

This paper is organized as follows. Section 2 explicates an identification task along with subjects, materials, and procedures and reports results. In section 3, Korean speakers' identification accuracy and reaction times are discussed and conclusion follows in section 4.

2. Experiment: Perception Study

2.1 Subjects

Twenty eight Korean learners of L2 English participated in the perceptual categorization, i.e., identification experiment. They were recruited from

³ As pointed out by an anonymous reviewer, it seems to be unclear whether English /l-r/ contrast can be categorized into the CG or the SC for Korean speakers. Actually, the data for /l-r/ identification was not included in Park and de Jong (2008) or de Jong and Cho (2012). However, from Jun's (2003) perception study showed that the perception of English /l/ was better than that of English /r/ for Korean speakers. And it is known that English /l/ is phonologically more similar to Korean /l/ than English retroflex /r/. These might lead us to speculate that English /l/ might be perceptually more similar to Korean /l/ rather than English /r/, and thus it is assumed that English /l-r/ contrast is categorized as the CG category.

campus community, and their major or minor was English. Twenty were females and eight males. They ranged in age from 20 to 26 with an average 22. The average period of formal English learning was 10 years, ranging from 6 to 16. Their self-rating proficiency of English was on average 5 out of 10 points in scale. All were native speakers of Korean and none reported any hearing impairment.

2.2 Stimuli

In order to obtain perception materials, one native speaker of English was asked to produce 140 nonce words written in English alphabet and 160 real English words embedded in a carrier sentence “Please say _____”. The stimuli were divided into nonce words and real words to see if there is any wordhood effect on the identification accuracy for all the contrasts.

The first type of stimuli was minimal pairs consisting of nonce words. They comprised 14 types of contrasts. 11 contrasts out of them were consonantal contrasts (/p-b/, /t-d/, /k-g/, /f-v/, /s-z/, /tʃ-dʒ/, /l-r/, /s-θ/, /d-ð/, /p-f/, /b-v/). Each token was a monosyllable containing an onset /z/, a nucleus (/i, u, e, o, aɪ/), and a coda (/p, b, t, d, k, g, f, v, s, z, tʃ, dʒ, l, r, s, θ, ð, v/). That is, the contrasting target consonants were placed in word-final position (e.g., *zeep, zeeb, zeet, zeed*...). 3 contrasts contained tense-lax vowel contrast (/i/-/ɪ/, /e/-/æ/, /u/-/ʊ/) (e.g., *zee, zit, zet-zat, zoo-zook*). In total, 140 nonce words were obtained for the identification task. See the full list of nonce words in Appendix.

The second type of minimal pairs was extracted randomly from real words which consisted of 16 contrasts. Similar to the nonce word set, 9 word sets ended in target consonants (/p-b/, /t-d/, /k-g/, /f-v/, /s-z/, /tʃ-dʒ/, /l-r/, /s-θ/, /d-ð/). Each token contained a target consonant in word-final position (e.g., *cap-cab, neat-need, buck-bug, piece-peas*...).⁴ 3 contrasts included the same sets of vowel contrast as the nonce word contrasts (/i/-/ɪ/, /e/-/æ/, /u/-/ʊ/) (e.g., *feel-fill, bed-bad, fool-full*...). Finally, 4 contrasts were obtained i.e., /p-f/, /b-v/, /s-/θ/, and /d-/ð/ in word-initial position. Unlike the pairs of a stop-a stop placed in word final position, these pairs consist of a stop and a fricative or a pair of fricatives with the different place of articulation. The rationale behind the selection of these pairs for word initial position is that they are expected to be rather difficult for some (Korean) L2 speakers to identify even in word initial position. For instance, according to Park and de Jong (2008: 714), English /p/ was labeled as /f/ second most frequently after /p/ and English /v/ as /b/ by 14%, next to as /v/ by 52%. Furthermore, it has

⁴ Since word-initial and syllable-initial positions are known to be marked rather than unmarked, comparatively more robust phonetic cues such as aspiration and duration are involved. Thus to avoid too much easiness of identification, more target consonant pairs were placed word-finally than word-initially. Note that as anonymous reviewers comment, it is expected that subjects might be sensitive to many acoustic cues relevant for voicing distinction, i.e., the duration of preceding vowel, the degree of released portion of the stops, etc.

long been observed that English interdentalals such as /θ, ð/ are substituted with /s/, /t/, or /d/ (Lombardi 2003). Given these previous studies, it is worth investigating these controversial pairs in the present study although the number of target pairs in word initial position is not exactly comparable to that in word final position. Hence, these four target pairs were placed in word-initial position (e.g., *pin-fin*, *bat-vat*, *sing-thing*, *day-they* ...). Since each set of contrast included 10 words, the total number of stimulus words was 160. See the full stimulus set in Appendix.

In sum, 300 tokens (140 nonce words and 160 real words) were presented to the subjects. Thus, a total of 8400 responses were obtained in the identification task and analyzed for statistics (300 tokens × 28 subjects).

2.3 Procedures

All of the perception stimuli were extracted from the recordings of the utterances produced by one native speaker of English. Before subjects participated in the perception task, they were first told in Korean about how they respond to each listening stimulus. The identification task was conducted in a quiet room on campus. The identification experiment proceeded in two blocks. In the first block, they sat in front of a computer and were asked to listen to a nonce word very carefully. The choices were visually presented simultaneously on the computer screen. Then, they were instructed to choose the word they heard and to press the button 1 or 2 on the keyboard as fast as possible (e.g., ① zeeb ② zeep). All the choices were visually presented in English orthography on the computer screen.

The second block proceeded in the same procedure as in the first block. Subjects heard real words and had to identify the word they heard as quickly as possible. They were asked to look at the two choices on the computer screen, choose one, and press the button 1 or 2 (e.g., ① cap ② cab).

All of the stimuli were visually presented and randomized by using the software E-Prime Professional 2.0. Synchronization of each token of listening stimulus and visual choices were controlled by running E-Prime. A word "Next" was placed with a pause of 2000ms after stimulus offset. Each identification task started after 5 trials. The entire experiment lasted approximately 20 minutes on average.

Since the main purpose of this research is to assess the identification patterns and accuracy for each L2 English contrast, accuracy and reaction times were measured through E-Prime. Note that if L2 listeners identify the segments or contrast relatively easily, it should take faster reaction time (Strange and Shafer 2008: 160). These two parameters, i.e., accuracy and reaction times were evaluated within each block and were calculated for each of the twenty eight subjects.

Accuracy and reaction times were compared by the potential factors such as L2 contrast type, voicing of consonants, manner of articulation, vowel type. Two parameters were repeatedly obtained within subjects, and they

were subjected to a one-way or two-way ANOVA. All the analyses were performed by IBM SPSS Statistics 20. The results of analyses are described as significant if $p < .05$ and highly significant if $p < .01$ on the basis of significance level $\alpha = 5\%$.

2.4 Results

2.4.1 Identification of L2 contrasts for real words

Table 1 shows the mean percent of identification accuracy and reaction time for three types of L2 category adopted under PAM. As predicted in the PAM model, the Korean speakers were found to encounter greater confusion distinguishing L2 English sound pairs classified as Single Category (57%) than Two Categories or Category Goodness. TC contrasts were most accurately identified (92%), and followed by CG contrasts (78%). Post-hoc pairwise comparison LSD analysis and Scheffe analysis displayed that three types of contrasts were significantly different from one another (all $p < .05$).

Furthermore, as shown in Table 1, the degree of difficulty of the identification based on accuracy is strongly corroborated by the results of reaction times. Korean speakers responded faster for the TC and CG contrasts than the SC contrast. A post-hoc Scheffe test revealed that RTs for TC and CG were significantly different from those for SG ($p < .05$).

The results both for accuracy and RTs indicate that the categorization of L2 sounds under the PAM model accounts for Korean speakers' difficulty of distinguishing L2 English contrasts. As expected, since L2 sound pairs such as /p-b/, and /s-z/ that belong to TC category can be perceptually mapped onto L1 Korean pairs such as /p^h-p/ and /s-tʃ/ respectively, they are highly likely to be distinguished. On the other hand, the L2 contrast such as /i-ɪ/ can be presumably assimilated to L1 Korean /i/ and thus more difficulty is predicted concerning the distinction.

Table 1. Identification accuracy and reaction time by category type

| | Category type | | | F & p-values |
|--------------|---------------|-----|-----|-------------------------------|
| | TC | CG | SC | |
| Accuracy (%) | 92 | 78 | 57 | $F[2,157]=38.56, p=.000^{**}$ |
| RT (sec.) | 2.1 | 2.2 | 2.8 | $F[2,157]=57, p=.000^{**}$ |

Another contributing goal of the present study was to reveal other potential factors which might explain the degree of difficulty of L2 speakers' identification of L2 contrasts. Thus an ANOVA was performed to see if there are any complementary factors to account for degrees of difficulty in identifying L2 sound contrasts. First, analysis displayed that consonant contrasts in English are better identified than vowel contrasts. To be specific, identification accuracy was much higher for consonantal contrast than for

vocalic contrast (85% vs. 57%, $F[1,157]=52.35$, $p=.000$). Furthermore, RTs were shorter for the former than for the latter (2.1 sec. vs. 2.8 sec., $F[1,157]=106.5$, $p=.000$). This finding indicates that Korean listeners identified English L2 contrasts for consonants more accurately and quickly than for vowels.

Additionally, a one-way ANOVA was performed to see if there are different degrees of identification problem according to manner of consonant articulation. The analysis showed that manner of articulation played a significant role in the identification. To be specific, accuracy was much higher for stop contrasts (/p-b/, /t-d/, /k-g/, 95%) and affricate contrast (/tʃ-dʒ/, 92%) than for fricative (/f-v/, /s-z/, /s-θ/, /d-ð/, /p-f/, 82%) and approximant contrasts (/l-r/, 84%) as shown in Table 2. This finding might be due to the fact that stop and affricate contrasts involved only TC, whereas fricative contrasts corresponded to both TC and CG category types. The accuracy and RT, however, did not differ between voiced and voiceless obstruents (e.g., /p-b/, $F[1,59]=.08$, $p>.05$), and interaction between voicing and manner did not reach significance ($F[2,59]=.11$, $p>.05$).

Table 2. Identification accuracy and reaction time by manner of articulation

| | Manner of articulation of target consonants | | | | F & p-values |
|--------------|---------------------------------------------|-------|---------|---------|--------------------------------|
| | Stop | Fric. | Affric. | Approx. | |
| Accuracy (%) | 95 | 82 | 92 | 84 | $F[3,89]=4.87$, $p=.004^{**}$ |
| RT (sec.) | 2.1 | 2.2 | 2.0 | 2.2 | $F[3,89]=1.21$, $p>.05$ |

Furthermore, we explored the possibility that there is any difference in the difficulty encountered by Korean speakers in identifying L2 English contrasts that belong to the identical TC category. The analysis showed that individual contrast type also exerted a crucial influence on the differences in identification accuracy ($F[5,59]=3.05$, $p=.01$) and RT ($F[5,59]=3.32$, $p=.01$). As plainly illustrated in Figure (1a), /s-z/ minimal pair was least accurately identified (77%) than the other pairs (all pairs over 90%) as revealed by post-hoc LSD and Duncan tests. With respect to RTs, two minimal pairs such as /p-b/ and /s-z/ were responded to rather slowly than the other pairs (approximately 2.3 sec for /p-b/ and /s-z/; 1.9~2.0 sec. for the other pairs) as illustrated in Figure (1b). This finding is of importance in that it indicates the major categorization proposed by the PAM model is not sufficient to account for the minor variation in difficulty of identification within the identical TC contrasts. In any case, it is not clear at this point why /s-z/ is peculiarly more difficult to identify than the other TC pairs for Korean speakers.

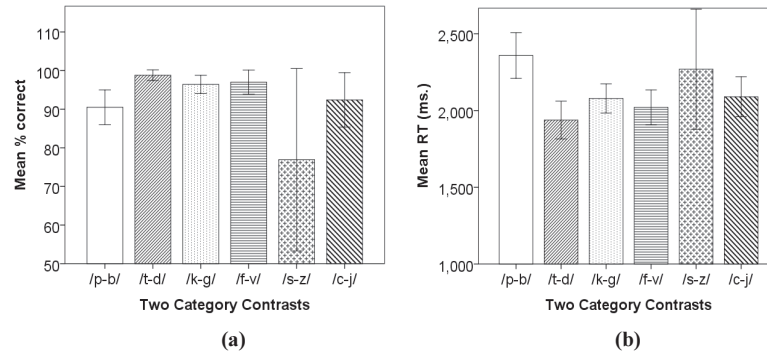


Figure 1. Mean % of correct responses (a), and mean RT (b) for the TC contrast

Next, we analyzed the data on the CG contrast to see if there is any difference in the degree of difficulty in identification in L2 contrasts within CG type. As is evident in Table 3, significant differences in identification accuracy and RT were found within the identical CG category minimal pairs. Interestingly, /s-θ/ pair was most accurately and rapidly identified whether it was in word-initial or final position. Furthermore, Korean listeners encountered the most difficulty in identifying the minimal pair /d-ð/ in terms of accuracy and RTs whether in word-initial or final position. A post-hoc Duncan test confirms that /d-ð/ sounds in word-initial and final position belong to the same group and the pairs /s-θ/ in word-initial and word-final position are grouped together. The pairs /l-r/ in word-final position, /p-f/, and /b-v/ in word-initial position are grouped together with similar accuracy. These findings indicate that there exists variation of difficulty in L2 contrast identification even within the CG type under the PAM model, suggesting inadequacy in predictability of the model. Under the PAM model, since English L2 contrast /l-r/ can be perceptually assimilated to Korean L1 /l/ ‘ㄹ’ and English /d-ð/ to Korean /t/ ‘ㄷ’, both contrasts belong to Category Goodness type and are expected to show a similar degree of difficulty of identification. The results we obtained are not consistent with the predictions of the PAM. Simple categorization of L2 contrasts does not seem to account for this variation of difficulty, and it may lead us to consider the perceptual assimilation from the perspective of phonetic cues as well as the sheer phonological categorization.

Table 3. Identification accuracy and reaction time within CG type

| | Word-final position | | | Word-initial position | | | | F & p-values |
|-----------|-----------------------|-------------------------|-------------------------|-----------------------|---------------------|----------------------|------------------------|---------------------------|
| | /l-r/ | /s-θ/ | /d-ð/ | /p-f/ | /b-v/ | /d-ð/ | /s-θ/ | |
| Ex. | <i>deal- deer</i> | <i>tense- tenth</i> | <i>seed- seethe</i> | <i>pin- fin</i> | <i>bet- vet</i> | <i>den- then</i> | <i>sing- thing</i> | |
| Acc. (%) | 84 | 90 | 55 | 89 | 75 | 63 | 95 | F[6,67]=10.1, p=.000** |
| RT (sec.) | 2.3 | 2.2 | 2.6 | 1.9 | 2.4 | 2.4 | 1.9 | F[6,67]=11.71, p=.000* |

The identification accuracy was examined for individual phonemes within each pair in CG type as shown in Figure 2. Among the pairs targeted in word-final position, /l/ was better identified than /r/ (90% vs. 77%), and /d/ faced the most difficulty in being distinguished from /ð/ (32% vs. 78%). The pair /d-ð/ was least accurately identified among the pairs in word-initial position (64% vs. 63%). It is not clear why Korean listeners face most confusion in distinguishing /d/ from /ð/ both in word-initial and word-final position. One possible explanation is that less marked consonants like a fricative are more difficult in perception and production than more marked ones like a stop (Eckman and Iverson 1993, Lombardi 2003, Jun 2003). Another suggestion is that these two sounds may share greater degree of perceived similarity than the other pairs. Presumably the weak noise energy of dental fricatives and lack of dental fricatives in Korean L1 phoneme inventory may contribute to such perceptual illusion (Ladefoged 2006: 194).

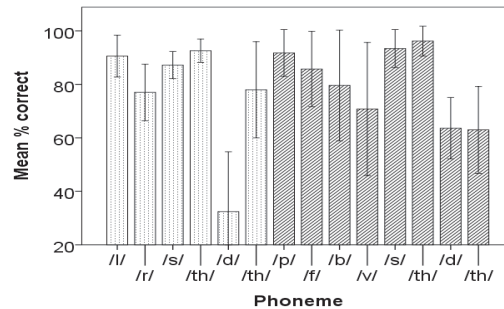


Figure 2. Mean % of correct responses by individual phonemes (dotted bars represent the minimal pairs in word-final position and dashed bars represent the pairs in word-initial position.)

Finally, we examined the possibility of differences in identification accuracy within SC condition. As mentioned earlier, as predicted by the PAM model, the SC contrast was more difficult to identify than the TC or the

CG contrasts.⁵ Analysis exhibited that each type of vowel pair within the SC contrast did not have an effect on the level of difficulty as clearly manifested in Table 4. Significant differences in both accuracy and RTs were not found. Furthermore, there were no effects of individual phonemes ($F[5,29]=.59$, $p>.05$), and vowel height ($F[2,29]=.4$, $p>.05$) on the identification accuracy as well as RTs. This finding seems to suggest that the prediction concerning the level of perceptual assimilation hold true at least for L2 English vowel contrasts since all the vowel pairs within the SC contrast show similar degree of difficulty of making a distinction between L2 contrasts.

Table 4. Identification accuracy and reaction time by vowel pair type within SC

| | Vowel pairs | | | F & p-values |
|--------------|---------------------------|-------------------------|---------------------------|-------------------------|
| | /i-ɪ/ <i>feel-fill</i> | /e-æ/ <i>bed-bad</i> | /u-ʊ/ <i>fool-full</i> | |
| Accuracy (%) | 59 | 55 | 52 | $F[2,29]=.4$, $p>.05$ |
| RT (sec.) | 2.8 | 2.7 | 3.1 | $F[2,29]=2.6$, $p>.05$ |

In summary, the results found for real words indicate that Korean L2 speakers identified L2 English sounds belonging to the TC contrast most accurately and rapidly, followed by those of the CG contrasts and then encountered most difficulty in distinguishing sounds within the SC contrast. These results are consistent with the prediction of the PAM model. However, we found the variation in the identification accuracy within the TC and CG contrasts. This might lead us to consider other factors such as the position of the consonantal contrasts in word (initial vs. final), manner of articulation of the consonants, or the acoustic properties of the individual target segments.

2.4.2 Identification of L2 contrasts for nonce words

As aforementioned, in order to eliminate the wordhood effect on the perceptual capacity for L2 sounds, analysis was conducted for nonce words. Table 5 shows the mean accuracy and reaction times by the category type under the PAM model (Two Category, Category Goodness, and Single Category). As plainly shown, the pattern of results for nonce words is consistent with that for real words reported in previous section. To be specific, the TC contrast was most accurately and rapidly identified, followed by the CG contrast, and lastly by the SC contrast. This pattern is exactly in conformity with the prediction under the PAM model. A post-hoc pairwise comparison analysis (LSD) displayed that there was no significant

⁵ As pointed out by an anonymous reviewer, the SC category applies only to vowel pairs, while the TC and CG contrasts hold true for consonant categories just as designed in the current experiment. Whether the L2 targets are consonants or vowels, the PAM model classifies all the contrasts depending of the degree of similarity or discrepancy between L1 and L2 sounds. Assuming that, the results in the current study are reported in the same fashion.

difference in the identification accuracy between TC and CG, but significant difference was found between CG/TC and SC ($p < .05$). What is different from the results for real words is that the CG contrast was better distinguished for nonce words than for real words (85% vs. 78%).

Table 5. Identification accuracy and reaction time by category type

| | Category type | | | F & p-values |
|--------------|---------------|-----|-----|-------------------------------|
| | TC | CG | SC | |
| Accuracy (%) | 88 | 85 | 53 | $F[2,139]=69.75, p=.000^{**}$ |
| RT (sec.) | 2.0 | 2.1 | 2.3 | $F[2,139]=13.31, p=.000^{**}$ |

Furthermore, analysis was carried out to see whether other factors such as vocalicity affect the L2 contrast. The analysis exhibited that consonant target contrasts were more accurately identified approximately by 34% than vocalic contrasts (87% vs. 53%). Furthermore, the former was responded more rapidly than the latter (2.0 sec. vs. 2.4 sec.).

Next, we examined whether manner of articulation of the target consonants in contrast has an effect on the identification level. An ANOVA showed that stop and affricate contrasts (e.g., *zeep-zeeb*, *zeetch-zeedge*) were significantly more accurately identified than fricative or approximant contrasts (e.g., *zeef-zeev*, *zeel-zeer*). This pattern as well is consistent with that found for real words. Results for both nonce and real words indicate that manner of articulation of the consonants play a crucial role in identifying L2 contrast in addition to category type.

Table 6. Identification accuracy and reaction time by manner of articulation

| | Manner of articulation of target consonants | | | | F & p-values |
|--------------|---------------------------------------------|-------|---------|---------|-----------------------------|
| | Stop | Fric. | Affric. | Approx. | |
| Accuracy (%) | 92 | 83 | 94 | 78 | $F[3,79]=8.85, p=.000^{**}$ |
| RT (sec.) | 2.0 | 2.1 | 2.0 | 2.1 | $F[3,79]=.59, p>.05$ |

Then let us turn to the potential difference in the ability to perceptually categorize L2 contrast within each category type. As seen in Figure 3(a), the identifiability of minimal pairs within the TC contrast showed variation ($F[5,59]=14.69, p=.000$). To be specific, the minimal pairs such as /t-d/, /k-g/ and /tʃ-dʒ/ were most accurately distinguished (respectively, by 97%, 95%, 94%). The pair /s-z/ in nonce words was found to be most difficult to identify (77%).⁶ Furthermore, the former was responded significantly faster than the latter (2.0 sec. vs. 2.1 sec., $F[5,59]=2.43, p=.04$) as shown in Figure 3(b). However, there was no voicing effect on the identification accuracy for the

⁶ The identification result found of /s-z/ pair in this study is very similar to that reported in Pilus (2005), where English /t-d/ pair was identified by 92.5% and /s-z/ pair was identified by 79% in word-final position by Malayan speakers of L2 English.

individual segments within each pair (e.g., *zeep-zeeb*, $F[1,59]=.09$, $p>.05$). These findings are consistent with those obtained for real words, suggesting that other factors such as individual contrast within the same TC contrast should be taken into account to predict the perceptual categorization for L2 contrasts.

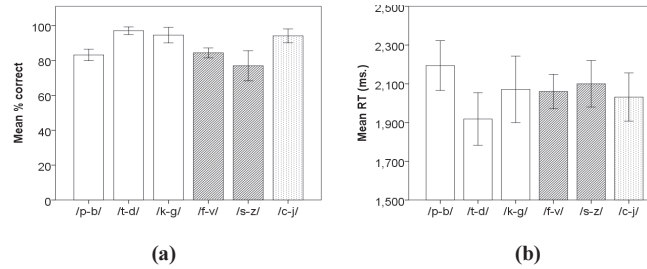


Figure 3. Mean % of correct responses (a), and mean RT (b) for the TC contrast

Of course, as might be easily guessed from Figure 4, it was found that there were significant differences in the identification accuracy according to individual phonemes ($F[11,59]=7.25$, $p=.000$) and in RTs as well ($F[11,59]=2.44$, $p=.01$). /t/, /g/, /dʒ/, /d/, /k/, and /tʃ/ were more accurately identified (above 90%) than the other phonemes. /z/ was found to be most difficult to identify (74%).

We also examined whether there is any difference in the discrimination ability for minimal pairs within the CG contrast. As is clearly seen in Table 7, there were significant differences in the identification accuracy. To be specific, /d-ð/ pair was most accurately identified whereas /l-r/ pair was least accurately identified. This pattern differs from that of the results found for real words reported in previous section. Note that, /d-ð/ pair was most difficult in word-final position in the real word targets (55%). It seems that this contrast is affected by the wordhood of the stimulus even though it is not clear why. Furthermore, post-hoc LSD and Duncan tests showed that /d-ð/ and /s-θ/ are grouped together and /l-r/, /p-f/ and /b-v/ belong to the same group. This finding suggests that the degree of difficulty predicted under the PAM model is not sophisticated to explain the variation within the same contrast type.

Table 7. Identification accuracy and reaction time within the CG contrast

| | /l-r/ | /s-θ/ | /d-ð/ | /p-f/ | /b-v/ | Mean | F & p-values |
|-----------|----------------------------------|-------------------------|-----------------------|-----------------------|-------|------|----------------------------|
| Ex. | <i>zeel-zeerzeace- zeeth</i> | <i>zeed- zeethe</i> | <i>zeep- zeef</i> | <i>zeeb- zeev</i> | | | |
| Acc. (%) | 78 | 87 | 92 | 84 | 84 | 85 | $F[4,49]=3.41$, $p=.01^*$ |
| RT (sec.) | 2.1 | 2.1 | 2.0 | 2.1 | 2.1 | 2.1 | $F[4,49]=.8$, $p>.05$ |

Looking into the accuracy of identification of individual consonants in word-final position, significant differences were found ($F[9,49]=7.49$, $p=.000^{**}$), but there was no difference in RTs ($F[9,49]=.75$, $p>.05$). As plainly illustrated in Figure 4, /θ, d, ð, r/ were relatively better identified than the other phonemes.

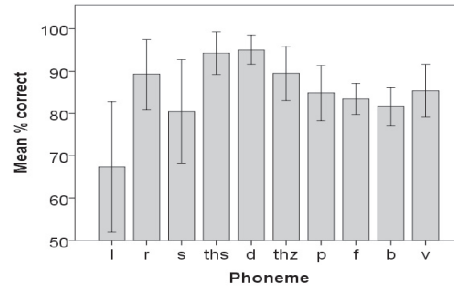


Figure 4. Mean % of correct responses by individual phonemes with the CG contrast

Finally, we looked into the SC contrast to see if there is any difference even within the most difficult category group. As demonstrated in Table 8, analysis exhibited that there was no significant differences with the vowel contrasts both in accuracy and in RTs. This finding is consistent with that obtained for real words. To be specific, as predicted by the PAM model, the minimal pairs in the SC contrast were most difficult to distinguish among all the contrast (50-60%).

Table 8. Identification accuracy and reaction time by vowel pair type within SC

| | Vowel pairs | | | F & p-values |
|--------------|---------------------------|-------------------------|---------------------------|--------------------------|
| | /i-ɪ/ <i>feel-fill</i> | /ɛ-æ/ <i>bed-bad</i> | /u-ʊ/ <i>fool-full</i> | |
| Accuracy (%) | 56 | 53 | 50 | $F[2,29]=.15$, $p>.05$ |
| RT (sec.) | 2.4 | 2.4 | 2.2 | $F[2,29]=1.36$, $p>.05$ |

Despite of uniform accuracy within the SC contrast, the identification accuracy exhibited the variation among individual vowels as illustrated in Figure 6 ($F[5,29]=9.78$, $p=.000^{**}$). Post-hoc LSD analysis displayed that vowel phonemes such as /ɪ, æ, ʊ/ were more accurately identified (respectively by 76%, 60%, 76%) than /i, ɛ, u/ (respectively, 37%, 45%, 24%). Significant differences were not found for RTs. Furthermore, there was no effect of vowel height (/i/ vs. /ɛ/ vs. /æ/). However, post-hoc LSD analysis revealed that identification accuracy was higher for lax vowels such as /ɪ, ʊ/ than for tense ones /i, u/ (all $p<.05$). Additionally, phonetically longer /æ/ was better identified than shorter /ɛ/ (60% vs. 46%) though the difference was not significant.

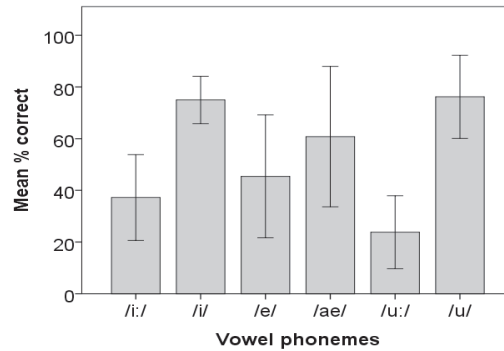


Figure 5. Mean % of correct responses by individual phonemes with the SC contrast

In summary, when Korean speakers heard L2 English sound contrasts embedded in nonce words, overall they showed a similar pattern in identification accuracy and RT to that found for real word condition. As predicted by the PAM model, minimal pairs belonging to the TC contrast were identified most accurately and rapidly, and then the CG contrast pairs showed relatively high accuracy, and the SC contrast group confronted with the most difficulty in being identified. However, different degrees of difficulty in perceptual assimilation were found even within the TC and CG contrasts, suggesting that the classification proposed by the PAM model is not sufficient and other factors should be taken into account. Our results showed that manner of articulation, the vocalicity of the segments, individual type of contrast, and the prosodic position of the contrast might be potential factors affecting the difficulty encountered by L2 speakers.

3. Discussion

3.1 Does the PAM model predict the Korean L2 speakers' relative difficulty in discriminating English sound contrasts?

The current study conducted a perception experiment with 16 pairs of L2 English contrasts carried in real words and with 14 pairs of similar contrasts embedded in nonce words to see if Korean listeners encounter relative difficulty in distinguishing voicing, manner, position-in-the word or relevant contrasts according to the prediction by the PAM model. The overall patterns that emerged from real word condition and nonce word condition were very similar in that the L2 identification accuracy was highest for the TC contrast (90%), intermediate for the CG contrast (82%), and lowest for the SC contrast (55%). Furthermore, the TC category pairs were responded more quickly than the other two category types (2.05 sec. vs. 2.35 sec.). This finding is exactly in accordance with the prediction by the PAM model. Furthermore, it is supported by Park and de Jong's (2008) study on the

mapping of English L2 sounds onto Korean counterparts in that the perceptual patterns of L2 pairs can be predicted, to some extent, by the phonological systems of L1 categories.

Since the L2 pairs within the TC contrast such as /p-b/, /t-d/, and /k-g/ can be perceptually assimilated to two Korean categories with almost the same degree of difficulty, the identification accuracy was highest and the RT was fastest.⁷ However, the minimal pairs belonging to the CG category such as /l-r/, /s-θ/, /d-ð/, etc. were encountered with lower accuracy and longer RTs by Korean speakers. This might be due to the fact that these two English sounds are perceptually assimilated to one Korean category with one member being more difficult to identify than the other. Finally, Korean listeners faced most difficulty in identifying the English minimal pairs within the SC such as /i-ɪ/, /ɛ-æ/ and /u-ʊ/. Since both English L2 vowels in each pair usually assimilate to one Korean vowel, respectively /i/, /e/, /u/ and both are perceptually equally distant from Korean counterparts, they are most difficult to identify. Such difficulty of identifying English lax-tense pairs can be seen from a different type of evidence shown by Mack's (2003) study. It showed that even Korean/English bilinguals who were exposed to English after the age of five failed to discriminate the boundary between /i/ and /ɪ/ unlike native speakers of English.

Overall, our results showed that the categorization type under the PAM model can be a predictor of the degree to which Korean listeners might encounter the relative difficulty in identifying English L2 contrast. However, it was found that L identification accuracy and/or reaction times exhibited variation even within a contrast type, especially the TC contrast and CG contrast. This finding cannot be offered an adequate account for under the PMA model. Thus other factors might contribute to the difficulty of various degrees of identification of L2 contrast, which is discussed in more detail in the next section.

3.2 How can we account for the various levels of difficulty of L2 contrast identification within the contrast type?

Looking in more detail at the identification accuracy and reaction time pattern with each of the category (TC, CG, SG), we found an interesting result that the identification accuracy differed by specific contrast, specifically within the TC and CG contrast. First, there were differences in the identification accuracy between the voicing contrast in stop/affricate pairs and the fricative pairs within the TC contrast. This might lead us to speculate that the acoustic cues for voicing for the former are more robust than those for the latter. For instance, the duration of the preceding vowel, the presence of aspiration or the amount of noise energy for stops and

⁷ As pointed out by an anonymous reviewer, it is plausible that the subtle acoustic properties such as the amount of release burst or the duration of preceding vowels due to the place of articulation can contribute to minor variation in the degree of identification of L2 pairs.

affricates might play greater cues for identifying voicing contrast than for fricatives. However, average differences in duration of the preceding vowels were not different between stop contrasts and fricative contrasts, especially in real words (139 ms vs. 141 ms). This suggests that the different degrees of identifiability between stops and fricatives cannot be attributable to a sole acoustic property but rather multiple cues might cooperate for the perception of the consonants. Furthermore, the differences in duration of the preceding vowel were greater for /t-d/ and /k-g/ contrasts, respectively 181 ms, and 130 ms, than for /p-b/ voicing contrast (105 ms). Such gaps might account for comparatively identification accuracy for the former than for the latter. However, a greater gap in the duration of the preceding vowel between /s/ and /z/ led to rather lower identification accuracy than that between /tʃ/ and /dʒ/. (Of course, note that duration of preceding vowels were significantly longer before voiced obstruents than voiceless ones, 241 ms vs. 111 ms, $F[1,58]=09.58$, $p=.000$.) Thus it seems that duration of the preceding vowel as well is not sufficient to account for the degree of identification accuracy. Thus a more comprehensive set of cues might await further investigation in future research. Hence one thing is clear: manner of articulation of the consonants can be considered as an additional side factor to distinguish L2 contrast, all being equal within the TC contrast.

Next, there were varying levels of difficulty in the identification accuracy with the CG contrast. As mentioned previously, the accuracy from the results for real word condition was highest for /s-θ/ both in word-initial and word-final position (90-95%), whereas the pair /d-ð/ was most difficult to identify in both the prosodic position (55-63%).⁸ This finding cannot be properly couched within PAM. One possible explanation is that voiceless pairs are more likely to be identified than their voiced counterparts probably because more robust acoustic cues are available for the former, leading to less confusion than the latter. Another explanation can be found in the SLM model (Flege 1995a,b). Under SLM, the more similar L2 sound to L1, the more difficult it is to acquire. De Jong and Cho (2012: 364) reported that English /θ/ is perceptually mapped onto Korean /s'/ by 48% and English /s/ onto Korean /s'/ by 95%. On the other hand, English /ð/ was perceived as Korean /t/ by 76%, and English /d/ as Korean /t/ by 95% (See similar percentages in Park and de Jong (2008: 709)). Putting these together, English /ð/ is more similar to Korean /t/ perceptually than English /θ/ to Korean /t/. Thus, the greater perceived similarity for the former is highly likely to cause greater difficulty in identifying /ð/. This line of interpretation might lead to the suggestion that both PAM and SLM should be taken into consideration to account for the identification accuracy pattern emerging between L2 sounds and between L2 and L1 sounds.

⁸ A production-based account does not give a suitable account for this asymmetry between these two pairs. It has been known that Korean learners of English tend to substitute /s/ for /θ/ and /d/ for /ð/ because Korean does not have a /z/ (Ioup 2008). However, such an account does not explain why /d-ð/ pair is more poorly identified than /s-θ/ pair.

Finally, we found that identification accuracy for /l-r/ pair was intermediate between stop voicing contrasts and fricative voicing contrasts both in real word and nonce word conditions (respectively, 84%, 78%). To be specific, /l-r/ pair was better identified than /f-v/ or /s-z/ pairs in real word condition alone. However, this pattern did not emerge in nonce word condition. Thus this inconsistency is just worth mentioning, but should be left for further comprehensive study. But in any case, the finding that one pair within the CG contrast is distinguished more accurately than another pair within the TC contrast cannot be captured by PAM.

4. Conclusion

The present research confirms that overall L2 identification accuracy and RT patterns are in accordance with the predictions of PAM. Furthermore, it shows that there are varying levels of identification patterns with the specific type of contrast, suggesting that complementary accounts should be provided to explain such variability. The possible alternatives include SLM to account for the relative difficulty to identify the L2 contrast by leaning on the perceived similarity between L2 and L1 sounds. Additionally, this study proposes that manner of articulation of consonants and vocalicity might also contribute to the relative difficulty of L2 identification. However, further study needs to be conducted to see whether there is any correlation between the acoustic properties of listening stimuli and perceptual patterns observed in the perception study. Or we need to probe into more comprehensive factors that might provide a systematic account for the varying levels of difficulty in identification tasks.

Appendix

Real word stimuli

| <i>Two Categories Stimuli</i> | | | | | |
|----------------------------------|--------------------------------|-------------|---------------|--------------|--------------|
| p-b | t-d | k-g | f-v | s-z | ʈʂ-dʒ |
| cap-cab | neat-need | buck-bug | loaf-loave | piece-peas | search-serge |
| cup-cub | coat-code | pick-pig | safe-save | loose-lose | rich-ridge |
| mop-mob | fate-fade | snack-snag | leaf-leave | ass-as | etch-edge |
| poop-poob | boot-boood | sock-sog | proof-prove | bus-buzz | batch-badge |
| sip-sib | not-not | peck-peg | strife-strive | price-prize | catch-cadge |
| <i>Category Goodness Stimuli</i> | | | | | |
| l-r | s-θ | d-ð | s-θ | d-ð | p-f |
| deer-deal | tense-tenth | seed-seethe | sing-thing | day-they | pin-fin |
| poor-pool | pass-path | ride-writhe | sick-thick | den-then | pat-fat |
| fear-feel | force-forth | sued-soothe | sigh-thigh | Dan-than | pail-fail |
| core-call | miss-myth | bade-bathe | sum-thumb | dough-though | pun-fun |
| tore-tall | face-faith | load-loath | seam-theme | dose-those | peel-feel |
| b-v | <i>Single Category Stimuli</i> | i-I | ɛ-æ | u-U | |
| bet-vet | | fill-feel | bed-bad | fool-full | |
| base-vase | | sit-seat | head-had | pool-pull | |
| bat-vat | | list-least | said-sad | suit-soot | |
| bolts-volts | | rich-reach | very-vary | Luke-look | |
| bail-veil | | ship-sheep | met-mat | wooed-would | |

Nonce word stimuli

| <i>Two category stimuli</i> | | | | | |
|----------------------------------|-------------|-------------|-----------|-------------|---------------|
| p-b | t-d | k-g | f-v | s-z | ʈʂ-dʒ |
| zeep-zeeb | zeet-zeed | zeek-zeeg | zeef-zeev | zeace-zeeze | zeetch-zeedge |
| zoop-zoob | zoot-zood | zook-zoog | zoof-zoov | zoose-zooze | zootch-zoodge |
| zape-zabe | zate-zade | zake-zague | zafe-zave | zase-zaze | zatch-zadge |
| zope-zobe | zote-zode | zoke-zogue | zofe-zove | zose-zoze | zotch-zidge |
| zipe-zibe | zite-zide | zike-zigue | zife-zive | zise-zize | zitch-zidge |
| <i>Category Goodness stimuli</i> | | | | | |
| l-r | s-θ | d-ð | p-f | b-v | |
| zeel-zeer | zeace-zeeth | zeed-zeethe | zeep-zeef | zeeb-zeev | |
| zool-zoor | zoos-zooth | zood-zooth | zoop-zoof | zoob-zoov | |
| zale-zare | zase-zath | zade-zathe | zape-zafe | zabe-zave | |
| zole-zore | zose-zoth | zode-zothe | zope-zofe | zobe-zove | |
| zile-zire | zise-zith | zode-zithe | zipe-zife | zibe-zive | |
| <i>Single Category stimuli</i> | | | | | |
| i-I | ɛ-æ | u-U | | | |
| zee-zit | zet-zat | zoo-zook | | | |
| kee-kit | ket-kat | koo-kook | | | |
| nee-nit | net-nat | noo-nook | | | |
| ree-rit | ret-rat | roo-rook | | | |
| chee-chit | chet-chat | choo-chook | | | |

REFERENCES

- BEST, CATHERINE. 1993. Emergence of language-specific constraints in perception of non-native speech: A window on early phonological development. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. MacNeilage and J. Morton (eds.). *Developmental Neurocognition: Speech and Face Processing in the First Year of Life*. Dordrecht: Kluwer Academic Publishers.
- _____. 1994. The emergence of native-language phonological influence in infants: A perceptual assimilation model. In C. Goodman and H. Nusbaum (eds.). *The Development of Speech Perception*, 167-224. Cambridge: The MIT Press.
- _____. 1995. A direct realist view of cross-language speech. In W. Strange (ed.). *Speech perception and linguistic experience*, 171-204. Baltimore: York Press.
- BEST, CATHERINE, GERALD MCROBERTS and NOMATHEMBA SITHOLE. 1988. The phonological basis of perceptual loss for non-native contrasts: Maintenance of discrimination among Zulu clicks by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance* 14, 345-360.
- DE JONG, KENNETH and MI-HUI CHO. 2012. Loanword phonology and perceptual mapping: Comparing two corpora of Korean contact with English. *Language* 88.2, 341-368.
- ECKMAN, FRED and GREGORY IVERSON. 1993. Sonority and markedness among onset clusters in the interlanguage of ESL learners. *Second Language Research* 9, 234-252.
- ESCUDEO, PAOLA, ELLEN SIMON and HOLGER MITTERER. 2012. The perception of English front vowels by North Holland and Flemish listeners: Acoustic similarity predicts and explains cross-linguistic and L2 perception. *Journal of Phonetics* 40, 280-288.
- FLEGE, JAMES. 1987. The production of new and similar phones in a foreign language-Evidence for the effect of equivalence classification. *Journal of Phonetics* 15.1, 47-65.
- _____. 1995a. Oddity discrimination of English vowels by non-natives. Manuscript prepared in the Department of Biocommunication, University of Alabama at Birmingham.
- _____. 1995b. Second language speech learning theory, findings, and problems. In W. Strange (ed.). *Speech perception and linguistic experience*, 233-277. Baltimore: York Press.

- _____. 2002. Interactions between the native and second-language phonetic systems. In P. Burmeister, T. Piske and A. Rohde (eds.). *An Integrated View of Language Development: Papers in Honor of Henning Wode*, 217-244. Trier: Wissenschaftlicher Verlag.
- FLEGE, JAMES and OCKE-SCHWEN BOHN. 1989. The perception of English vowels by native Spanish speakers. *Journal of the Acoustical Society of America* 85:S85(A).
- FLEGE, JAMES, OCKE-SCHWEN BOHN and S. JANG. 1997. The effect of experience on nonnative subjects' production and perception of English vowels. *Journal of Phonetics* 25, 169-186.
- FLEGE, JAMES and WIEKE EEFING. 1987. The production and perception of English stops by Spanish speakers of English. *Journal of Phonetics* 15, 67-83.
- FLEGE, JAMES and MURRAY MUNRO. 1994. The word unit in L2 speech production and perception. *Studies in Second Language Acquisition* 16, 381-411.
- GILICHINSKAYA, YANA and WINIFRED STRANGE. 2010. Perceptual assimilation of American English vowel by native Russian listeners. *Journal of the Acoustical Society of America* 128, EL80-EL85.
- IOUP, GEORGETTE. 2008. Exploring the role of age in L1 phonology. In J. G. Edwards and M. Zampini (eds.). *Phonology and Second Language Acquisition*, 41-62. Amsterdam/Philadelphia: John Benjamins Publication.
- JAMES, ALLAN. 1984. Phonetic transfer: The structural bases of interlinguals assessment. In M. Van den Broecke and A. Cohen (eds.). *Proceedings of the Tenth International Congress of Phonetic Sciences*. Dordrecht, The Netherlands: Foris.
- JUN, EUN. 2003. The perception of English /l/ and /r/ by Korean speakers. *Studies in Phonetics, Phonology and Morphology* 9.2, 471-484.
- LADEFOGED, PETER. 2006. *A Course in Phonetics* (Fifth Edition). Thomson Wadsworth.
- LOMBARDI, LINDA. 2003. Second language data and constraints on manner: Explaining substitutions for the English interdental. *Second Language Research* 19.3, 225-250.
- MACK, MOLLY. 2003. The phonetic systems of bilinguals. In M. Banich and M. Mack (eds.). *Mind, Brain, and Language*, Mahwah, NJ: Lawrence Erlbaum Associates, 309-349.
- MAJOR, ROY. 1987. Phonological similarity, markedness, and rate of L2

- acquisition. *Studies in Second Language Acquisition* 9, 63-82.
- PARK, HANYONG and KENNETH de JONG. 2008. Perceptual category mapping between English and Korean prevocalic obstruents: Evidence from mapping effects in second language identification skills. *Journal of Phonetics* 36.4, 704-723.
- PILUS, ZAHARIAH. 2005. Perception of voicing in English-word-final obstruents by Malay speakers of English: Examining the Perceptual Assimilation Model. *Malaysian Journal of ELT Research*, 1-12.
- ROCHET, BERNARD. 1995. Perception and production of second-language speech sounds by adults. In W. Strange (ed.). *Speech Perception and Linguistic Experience: Issues in Cross-language Speech Research*, 379-410. Timonium, MD: New York Press.
- SEKIYAMA, KAORU and YOH'ICHI TOHKURA. 1993. Inter-language differences in the influences of visual cues in speech perception. *Journal of Phonetics* 21, 427-444.
- SON, GAYEON. 2008. *Production and Perception of English Vowel Categories by Native Korean Speakers*. MA Thesis. Ewha Womans University.
- STRANGE, WINIFRED, OCKE-SCHWEN BOHN, K. NISHI and S. A. TRENT. 2005. Contextual variation in the acoustic and perceptual similarity of North German and American English vowels. *Journal of the Acoustical Society of America* 118, 1751-1762.
- STRANGE, WINIFRED and VALERIE SHAFER. 2008. Speech perception in second language learners. In Jette Edwards and Mary Zampini (eds.). *Phonology and Second Language Acquisition*, 153-192. John Benjamins.
- TAKAGI, NAOYUKI. 1993. *Perception of American English /r/ and /l/ by Adult Japanese Learners of English: A Unified View*. PhD Dissertation. University of California-Irvine.

Gwanhi Yun
Department of English Language and Literature
Daegu University
201 Daegudaero, Jillyang, Gyeongsan, Gyeongbuk
Korea 712-714
e-mail: ghyun@daegu.ac.kr

received: March 10, 2014
revised: June 10, 2014
accepted: July 28, 2014