

A longitudinal study of individual differences in the acquisition of new vowel contrasts

Donghyun Kim^{a,*}, Meghan Clayards^{a,b}, Heather Goad^a

*Corresponding author: donghyun.kim@mail.mcgill.ca

^a Department of Linguistics, McGill University, 1085 Dr. Penfield, Montreal, QC, H3A 1A7, Canada

^b School of Communication Sciences and Disorders, McGill University, McGill College 8th Floor, Montreal, QC, H3A 1G1, Canada

Abstract

This study explores how individuals' second language cue weighting strategies change over time and across different contrasts. The study investigates the developmental changes in perceptual cue weighting of two English vowel contrasts (/i/-/ɪ/ and /ɛ/-/æ/) by adult and child Korean learners of English during their first year of immersion in Canada. Longitudinal results revealed that adult learners had an initial advantage in L2 perceptual acquisition over children at least for the /i/-/ɪ/ contrast, but after one year some children showed greater improvements especially on the more difficult /ɛ/-/æ/ contrast. Both groups of Korean learners showed different acquisition patterns between the two vowel contrasts: they used both spectral and duration cues to distinguish /i/-/ɪ/ but generally only duration to distinguish /ɛ/-/æ/. By examining cue weights over time, this study partially confirmed the hypothesized developmental stages for the acquisition of L2 vowels first proposed by Escudero (2000) for Spanish learners of English. However, some unpredicted patterns were also identified. Most importantly, the longitudinal results suggest that individual differences in cue weighting are not merely random variability in the learner's response patterns, but are systematically associated with the developmental trajectories of individual learners and those trajectories vary according to vowel contrast.

Keywords

L2 speech perception, Individual differences, Cue weighting, Longitudinal study, Vowel contrasts, Korean, English

A longitudinal study of individual differences in the acquisition of new vowel contrasts

1.0 Introduction

For a given language and sound contrast, listeners pay more attention to some acoustic dimensions over others and the relative importance of these cues in determining category identity is referred to as *cue weighting* (e.g., Francis, Kaganovich, & Driscoll-Huber, 2008; Holt & Lotto, 2006). During development, the acquisition of appropriate acoustic cue weights is essential for creating target-like phonetic and phonological categories and subsequently comprehending speech (Mayo, Scobbie, Hewlett, & Waters, 2003; Mayo & Turk, 2005; Nittrouer & Lowenstein, 2010). In research on second language (L2) speech perception, it is well established that learners often have difficulty with L2 contrasts because they initially rely on different acoustic cues than native (L1) listeners do. For example, in differentiating English tense and lax vowel contrasts, native English listeners rely predominantly on spectral cues with vowel duration being secondary whereas Spanish L2 learners rely predominantly on vowel duration despite not having contrastive vowel length in their native language (Casillas, 2015; Escudero, 2000; Kondaurova & Francis, 2008, 2010; Morrison, 2008). Spanish learners of Dutch also rely more heavily on duration than spectral differences for the Dutch /a:/-/a/ contrast while native Dutch listeners rely most on spectral differences (Escudero, Benders, & Lipski, 2009; Lipski, Escudero, & Benders, 2012). Similarly, Japanese learners' difficulty in perceiving the English /l/-/ɹ/ contrast has been attributed to their attending mostly to differences in second formant frequencies while native listeners are most sensitive to differences in the third formant (Iverson, Kuhl, Akahane-Yamada, Diesch, Tohkura, Kettermann, & Siebert, 2003). Thus consistent mismatches have been identified in the use of acoustic cues by L2 listeners and L1 listeners. Of particular interest to the present study is how and to what extent these mismatches are resolved through the course of learning an L2. We take an individual differences approach using longitudinal data to investigate how individual L2 learners differ in their developmental trajectories over time.

1.1 Individual differences in L2 speech acquisition

The previously cited works report on group level observations, which may mask significant individual variability in L2 speech learning (cf. Amengual, 2016a for bilingual individuals' perception and production of vowel contrasts; Darcy, Park, & Yang, 2015 for the relation between individual cognitive abilities and L2 phonological processing; Lengeris, 2009 for individual differences in L2 vowel processing in relation to L1 vowel processing; Schertz, Cho, Lotto, & Warner, 2015, 2016 for individual differences in cue weighting strategies; Sebastián-Gallés & Baus, 2005 for individual L2 learners' performance across different perceptual tasks). It has commonly been observed that some learners demonstrate some degree of native-like pronunciation while others are judged as less intelligible due to their foreign accent (Ioup, 2008). In L2 speech perception, previous research has found various sources for individual differences, with learners showing differences in cognitive abilities (Darcy et al., 2015), in perceptual training outcomes for L2 contrasts (Golestani & Zatorre, 2009; Perrachione, Lee, Ha, & Wong, 2011), and in degrees of sensitivity to L1 phonetic contrasts (Díaz, Baus, Escera, Costa, & Sebastián-Gallés, 2008). Cross-sectional studies of L2 learners' cue weights have also revealed distinct patterns of individual variability (Escudero, 2000; Kong & Edwards, 2015; Schertz et al., 2015, 2016; Wanrooij, Escudero, & Raijmakers, 2013) with some groups of learners coming closer to native-like performance than others. Consequently, some researchers have proposed that these groups reflect different developmental stages in the acquisition of L2 contrasts (Escudero, 2000; Morrison, 2008) although developmental changes in the weighting of acoustic-phonetic cues do not always reflect degree of experience with the target language (Escudero, 2000). The particular emphasis on individual differences in the present study is on the kind of individual differences that exist among learners in their cue weighting strategies, and

how individuals differ in their developmental trajectories. We investigate this by tracking L2 perceptual cue weights for two English front vowel contrasts in two groups of Korean learners of English (adults and children) with very similar levels of exposure over the course of the first year of immersion in an L2-speaking environment.

1.2 Developmental trajectories of L2 speech acquisition

The majority of previous research in L2 speech acquisition has employed cross-sectional comparisons of different variables such as age, length of residence (LOR), and L2 use to investigate the developmental process (Flege & MacKay, 2004). Cross-sectional studies alone, however, might not suffice to provide clear insight into some aspects of the development of language over time. Most importantly, cross-sectional studies cannot tell us if individuals who behave differently from one another are progressing through different stages or following different developmental paths. In contrast, examining developmental trajectories and individual differences along the way could be highly informative because it could shed light on whether large differences in starting states and end states of acquisition are related. The present study complements and extends previous cross-sectional studies by adopting a longitudinal approach to examining developmental changes in perceptual cue weighting. It has been observed that the field of second language acquisition suffers from a lack of longitudinal research (Ortega & Iberri-Shea, 2005) even though L2 developmental processes can be best captured by longitudinal studies. Only a few previous studies have reported the longitudinal development of L2 speech learning, such as the perception and production of English /l/ and /ɹ/ by Japanese learners of English (Aoyama et al., 2004), the production of English vowels by Japanese learners of English (Oh, Guion-Anderson, Aoyama, Flege, Akahane-Yamada, & Yamada, 2011), and the perception and production of English vowels by native Koreans (Tsukada et al., 2005). The findings of Tsukada et al. (2005) showed that there were no significant effects of LOR on Korean speakers' performance on English vowel perception and production over a period of approximately one year. However, their study included Koreans who had already been in the US for more than two years at the onset of the study. It is likely that most of the participants in their study were no longer in the earliest stages of L2 acquisition when learning is presumably most rapid, and thus it may have been difficult to observe developmental changes and significant effects of LOR within a year. The present study was designed to address this by tracking recent arrivals during their first year of residence in an L2 environment. Thus, we study learners from the early stages of exposure to native speaker input in an English-speaking environment and track their developmental trajectories for a considerable period of time.

A handful of studies have hypothesized stage-like development in the perceptual weighting patterns of L2 learners (Escudero, 2000; Morrison, 2008). Escudero (2000) proposed four hypothetical stages in the development of a new /i/-/ɪ/ contrast by native Spanish learners of English. The proposed sequence is that, initially, naïve Spanish learners are not able to identify tokens of /i/ versus /ɪ/, thereby showing a no-contrast pattern. At the next stage, the contrast is distinguished exclusively by duration information present in the input. Then, learners use both duration and spectral information but still give priority to duration cues. At the final stage, learners show English-like use of both spectral and duration cues with primary weighting appropriately placed on spectral cues. Although a developmental sequence was hypothesized in Escudero's (2000) work, this stage-like development was inferred from cross-sectional patterns. The present study investigates whether L2 learners progress through specific developmental stages, as hypothesized by Escudero (2000), using longitudinal data. In the present study, we extend the hypothesized developmental stages—developed by Escudero (2000) for native Spanish listeners—to learners with a different native language background (i.e., Korean learners of English). We expect that Korean learners of English will show similar developmental patterns since both Korean and Spanish learners of English encounter similar difficulties in acquiring the tense and lax vowel contrasts in English (Flege et al., 1997) and both must acquire these new vowel contrasts from a smaller vowel inventory in their L1.

1.3 The target vowel contrasts

Another goal of the current paper is to compare individual cue weighting patterns across different speech sound contrasts. Studies of L2 cue weighting have thus far typically tested a single contrast within the same individuals. We will compare two related contrasts, which will allow us to test how systematic individual differences are and whether developmental trajectories vary across contrasts. The target contrasts are the English front vowel contrasts /i/-/ɪ/ and /ε/-/æ/. Previous work has shown that /i/-/ɪ/ and /ε/-/æ/ are spectrally distinct, and that /i/ and /æ/ tend to be longer than /ɪ/ and /ε/ respectively in the productions of native English speakers (Hillenbrand, Clark, & Houde, 2000; Hillenbrand, Getty, Clark, & Wheeler, 1995). These vowel pairs are considered to be among the most difficult contrasts for L2 learners of English to acquire (Bion, Escudero, Rauber, & Baptista, 2006; Flege, Bohn, & Jang, 1997; Ingram & Park, 1997; Kondaurova & Francis, 2008, 2010; Tsukada et al., 2005). For example, Flege et al. (1997) examined the perception and production of English front vowels by groups of German, Spanish, Mandarin, and Korean learners of English with varying LOR in the US. They found that native English controls produced large spectral differences but small duration differences between /i/ and /ɪ/ as well as between /ε/ and /æ/. In contrast, inexperienced non-native groups failed to produce spectral differences (i.e., Spanish, Mandarin, Korean participants for /i/-/ɪ/ and German, Mandarin, Korean participants for /ε/-/æ/) but instead produced exaggerated duration differences to distinguish the vowel contrasts (i.e., German, Mandarin, Korean participants for /i/-/ɪ/). In perception, participants were tested on synthetic stimuli with varying spectral and duration cues. The results also showed that native English listeners relied primarily on spectral cues to categorize the vowel contrasts, whereas inexperienced learners of English relied on duration cues (i.e., Spanish, Mandarin, Korean participants for /i/-/ɪ/ and German, Mandarin, Korean participants for /ε/-/æ/). These results were corroborated in subsequent studies of perception in Korean adult learners of English (Lee, 2009; Kim, 2010) as well as in studies of perception and production in Korean children and adults (Tsukada et al., 2005).

The sources of the difficulties that L2 learners of English have when perceiving and producing new vowel contrasts likely stem from differences between the vowel inventories of the learners' L1 and the target L2 (Tsukada et al., 2005). Regarding Korean learners of English, previous studies have shown that English /i/-/ɪ/ and /ε/-/æ/ were perceptually assimilated to a single Korean vowel, /i/ and /ε/, respectively (Ingram & Park, 1997; Son, 2008; Yun, 2014, but see Baker, Trofimovich, Mack, & Flege, 2002). These studies have suggested that this pattern of behavior falls under single category assimilation as proposed by the Perceptual Assimilation Model (PAM: Best, 1995; PAM-L2: Best & Tyler, 2007). They have also reported that although both /i/-/ɪ/ and /ε/-/æ/ are highly problematic for Korean learners of English, learners have more perceptual and production problems with /ε/-/æ/ than with /i/-/ɪ/. While some of this may be due in part to differences in how well each of these vowels are assimilated to L1 categories (an issue we return to in the discussion), questions remain about differences in difficulty between these two vowel contrasts. Following these earlier studies, we also expect that Korean learners will have more difficulty in learning appropriate cue weights for the /ε/-/æ/ contrast, but by examining cue weights over time, we can determine whether the two contrasts display different developmental trajectories or whether the development of the harder contrast simply progresses at a different rate than that of the easier one. This will contribute to our understanding of the difference in difficulty between these two contrasts.

1.4 The effect of age on L2 speech acquisition

The present study also examines differences between adults and children in the acquisition of L2 speech sound categories. Most research on L2 speech learning has focused on adult learners, and relatively little research has examined the effect of age on the development of L2 speech (cf. Aoyama, Flege, Guion-Anderson, Akahane-Yamada, & Yamada, 2004; Aoyama, Guion-Anderson, Flege, Yamada, & Akahane-Yamada, 2008; Oh et al., 2011; Tsukada et al., 2005). Previous phonetic training studies on adult-child differences in L2 speech, where the amount of exposure is controlled rather than naturalistic, have found mixed results. Some studies have

found similar amounts of learning for all age groups (Giannakopoulou, Brown, Clayards, & Wonnacott, 2017; Heeren & Schouten, 2010), while others have found greater learning for children than adults (Giannakopoulou, Uther, & Ylinen, 2013) or outperformance of older children over both younger children and adults (Shinohara & Iverson, 2013, 2015).

More relevant to the present study is research on the effect of age on L2 speech learning in a naturalistic setting. Among the few studies that have examined age-related differences in L2 speech learning in a naturalistic environment, Snow and Hoefnagel-Höhle (1977) suggested that adults and adolescents can be more successful than children in learning to pronounce L2 sounds up to 4-5 months of immersion exposure. However subsequent studies have tended to find the opposite. Tsukada et al. (2005) looked for differences in the perception and production of English vowels between native Korean adults and children with differing LOR in the US. The results of their study showed that the Korean children's English vowel perception and production were more accurate than those of the Korean adults, suggesting that L2 speech learning in children may take place more rapidly during the first years of immersion. Aoyama et al. (2004, 2008) investigated L2 speech learning by Japanese adults and children learning English during their first years of immersion in the US to examine whether adult and child learners differ in their learning patterns. Aoyama et al. (2004), which explored the perception and production of English /l/, /ɪ/, and /w/, showed that the Japanese children made greater improvements in their learning of L2 segments than the Japanese adults did over the course of one year. In their later study, Aoyama et al. (2008) examined L2 phonological acquisition of English fricatives in Japanese adults and children learning English and found that the adults in their study had an initial advantage over the children, likely due to their previous English education, but after one year of immersion, the children's production of English fricatives improved more than the adults did, leading to a loss of differences between the two populations of learners. Oh et al. (2011) also examined English vowel production by Japanese adults and children and found that the children outperformed the adults in vowel production and furthermore that the children's production results showed native-like accuracy within a year's time in an English-speaking environment. Similarly, more successful learning of L2 speech sounds by children was reported in Baker, Trofimovich, Flege, Mack, and Halter (2008). They found that Korean children outperformed Korean adults in their perception and production of English vowel contrasts. They accounted for adult-child differences in L2 speech learning in terms of the interaction between L1 and L2 sound categories where children are less likely to be affected by their L1 and thus can be more successful in L2 speech learning.

Taken together, these studies suggest that younger learners have an advantage in learning phonological aspects of an L2. Based on the findings in Aoyama et al. (2004, 2008), we predict that Korean adult learners may have an initial advantage over the Korean child learners because of the adults' greater previous experience with English. We also expect that L2 speech learning in children will occur more quickly than in adults over the course of the immersion period. Furthermore, insofar as children are more likely to reach native-like levels, children and adults may show different developmental patterns. For example, children may not progress through the same stages as adults, they may have qualitatively different starting points, they may be less likely to attend to the wrong cues, and they may be less affected by hard versus easy contrasts. Thus, including both adults and children in the present study may allow us to arrive at a more complete characterization of development.

1.5 The present study

The present study seeks to build on the previous research described above. Most studies of L2 acquisition of sound categories have focused either on comparing the learning of different contrasts (Best & Tyler, 2007; Flege, 1995) or on the learning of appropriate cue weights for particular contrasts (Escudero et al., 2009; Kondaurova & Francis, 2008, 2010). Most studies have also used cross-sectional approaches to investigate the development of L2 speech acquisition (Escudero, 2000). Moreover, relatively little has been reported on age-related differences in L2 speech acquisition (Aoyama et al., 2004; Tsukada et al., 2005). Taken together, this has limited our understanding of (1) how individuals' cue weighting strategies

change over time, (2) how cue weights are acquired across contrasts, especially between easier and harder contrasts that rely on the same acoustic cues such as /i/-/ɪ/ and /ɛ/-/æ/, and (3) whether adult and child learners show different developmental patterns. This study addresses these issues by examining the developmental changes in the weighting of acoustic-phonetic cues by Korean learners of English during their first year of immersion in Canada. More specifically, Korean adult and child learners' developmental changes in perceptual cue weighting in terms of spectral and duration cues are assessed over time, testing the hypothesized developmental sequence for the acquisition of non-native vowel contrasts proposed by Escudero (2000) longitudinally. We also test whether child and adult learners differ in their acquisition patterns and whether they follow different trajectories.

2.0 Methods

2.1 Participants

Twenty-seven native Korean learners of English, 12 adult (KA; $M = 39.6$ years, $SD = 3.2$, range = 35–45) and 15 children (KC; $M = 8.9$ years, $SD = 1.9$, range = 6–12), participated in this study upon their arrival in Canada. At the onset of the study, participants had resided in Canada approximately 2 months ($M = 2.1$ months, $SD = 1.5$, range = 1–5) and had never lived in an English-speaking country prior to their arrival. All adult participants (12 female) were the mothers of the child participants (6 female, 9 male). There were 12 families in total, and they can be identified by their participant number (e.g., KC03 = the child from family 03, KC05b = one of two children from family 05). The participants were tested at regular intervals over the course of a year: at 4 months (Time 2), 8 months (Time 3), and 12 months (Time 4) after the first experiment (Time 1). One of the child participants was dropped after completing the Time 1 experiment due to his return to Korea, so the data from this participant were not included in the analysis. One adult and one child (KA08 and KC08) were not available for Time 4, so 24 participants (11 adults and 13 children) participated in the experiment at Time 4. Participants' demographic and language background information is provided in the appendix (Table A1). Ten native listeners of Canadian English (NE; $M = 23.6$ years, $SD = 4.4$, range = 18–30) also participated in the experiment as controls.¹ In the present study, although child native listeners were not included as native controls, we assume that native English adult and child listeners weight acoustic cues in vowel categorization in similar ways based on previous research on age-related differences in cue weighting for vowels (Gerrits, 2001).

All of the Korean participants came to Canada solely for the purpose of learning English in an English-speaking country for a couple of years. Thus, they all can be regarded as highly motivated learners of English. Adult participants attended English language courses in language schools and child participants attended primary or secondary school in English beginning at Time 1 or shortly after. Adult participants attended English classes 5 days per week ($M = 4.5$ hours of study per day, $SD = 1.0$, range = 3–7), and child participants attended school 5 days a week and also studied English after school ($M = 6.2$ hours of study per day, $SD = 1.1$, range = 4–8). All 12 adult participants had studied English at school in Korea before arriving in Canada ($M = 8.9$ years, $SD = 2.0$, range = 6–12). Twelve child participants (all child participants except for KC09 and KC10) had studied English in Korea, but for fewer years than the adult participants ($M = 2.4$ years, $SD = 1.8$, range = 0–5).² In spite of this previous exposure to English, all of the Korean participants had very little or no interaction with native speakers of

¹ The Korean and native English participants recruited for this study were part of a larger project. Additional data collected at Time 4 (see below) is reported in Kim, Clayards, and Goad (2017). Some of the details about participants, stimuli and procedures are repeated here for completeness. The control group data (native English) is also repeated here for the same reason.

² For the majority of Korean learners, years of English education in Korea (YOE in Table A1 in the appendix) merely shows the number of years between the first and the last experience of formal English instruction in school.

English before they came to Canada. In fact, the English education that the Korean adults received in South Korea was based on the grammar-translation method (Li, 1998). Even today, the majority of formal English education in South Korea is based on this method. Thus, we assume that all learners were in the early stages of L2 speech acquisition with similar levels of exposure to authentic spoken language input by native speakers after they arrived in Canada. None of the participants reported speech or hearing impairments, and all participants were paid for their participation in the experiment at each time point.

2.2 Stimuli

Perceptual stimuli consisted of two continua varying orthogonally in both spectral quality and duration of the vowel. There were two steps involved in the construction of the continua. First, two resynthesized natural speech continua were created based on recordings of /bit/, /bɪt/, /bet/, and /bæt/ produced by a male native speaker of Canadian English in his early 20s. These continua were made using *TANDEM-STRAIGHT*, which is a high quality vocoder that allows for the creation of natural-sounding continua between two naturally-produced endpoints (Kawahara, Takahashi, Morise, & Banno, 2009). It should be noted that the sound quality of the spectral continua used in the present study represents an improvement over other methods of creating stimuli (e.g., *Klatt synthesis or LPC resynthesis*) as it stays close to the original natural speech by morphing holistically between two naturally-produced endpoints. Two thirty-step continua from *bit* to *beat* and from *bet* to *bat* were generated using *TANDEM-STRAIGHT*. These continua varied in both spectral quality and duration between the two natural endpoints. From these 30 steps for each continuum, seven acoustically and auditorily distinct stimuli along the continuum were chosen including the two natural end points for each vowel contrast. Second, each of the seven steps along the natural continuum was manipulated using the PSOLA algorithm in Praat (ver. 5.3.55, Boersma & Weenick, 2013) to create 7 step vowel duration continua ranging from 70 to 250 ms for *bit-beat* and from 90 to 270 ms for *bet-bat* (30 ms/step). Each step along the duration continuum was slightly over one just-noticeable difference for native English listeners, namely 25 ms as reported in Klatt (1976). Therefore, 49 tokens for *bit-beat* (7 spectral steps × 7 duration steps) and 49 tokens for *bet-bat* (7 spectral steps × 7 duration steps) were created. Four repetitions (196 trials) of each stimulus were played in random order to a separate group of six native English listeners in a preliminary test to check the validity of the stimuli and to check whether category boundaries between contrasts were well represented by steps in the continua. After confirming the validity of the stimuli for each vowel contrast, five spectral steps from the original 30 spectral steps for each vowel contrast were chosen including the two end points and the category boundary. Finally, a continuum ranging in vowel duration from 70 to 230 ms for *bit-beat* and from 100 to 260 ms for *bet-bat* were created (40 ms/step) from each of the 5 steps along the vowel spectral continuum, creating 25 tokens for each *bit-beat* and *bet-bat* continuum, respectively. We used 5 step continua rather than 7 step continua as in the preliminary test to limit the length of the test for the child participants. The first and second formant (F1, F2) values (in Hz) and the duration values (in ms) for each vowel continuum are provided in Table 1.

Table 1 F1 & F2 values and durations of the vowels in the /i/-/ɪ/ and /ɛ/-/æ/ continua.

Step	/i/-/ɪ/			/ɛ/-/æ/		
	F1 (Hz)	F2 (Hz)	Duration (ms)	F1 (Hz)	F2 (Hz)	Duration (ms)
1	437	1938	70	654	1710	100
2	366	2202	110	670	1684	140
3	333	2323	150	697	1677	180
4	312	2409	190	756	1674	220
5	269	2463	230	850	1704	260

2.3 Procedure

Cue weighting stimuli were presented using a two-alternative forced choice identification task created in *Praat*. Participants were tested one at a time in a quiet room immediately after completing a production task as part of a larger project. To avoid orthographic bias, pictures were used to represent the response words. Participants heard the stimulus once and pressed the left (←) or right (→) arrow on the keyboard to identify the picture that corresponds to the word they heard. Instructions were administered in Korean on the computer screen as well as orally. Participants were told that each trial was self-paced with no limit on time to respond. There were 25 stimuli repeated five times each in each of the two blocks for each vowel contrast (250 trials in total). Each block was counterbalanced, and all trials within a block were randomly presented through headphones at a comfortable listening level. All participants were asked to provide the words representing the pictures before the identification tests in order to ensure that they recognized what the pictures represented. Before the experiment began, there was a practice session consisting of up to ten trials with naturally produced tokens of the words *sheep* and *ship* to familiarize listeners with the procedure. Once the participants indicated that they understood the procedure, the practice session stopped before reaching the tenth trial to minimize training effects. The structure of the practice session was identical to the actual experiment, and the practice session was carried out only at Time 1.

3.0 Results

3.1 Group results

The participants' responses were analyzed using mixed-effects logistic regression and model comparisons to identify the fixed and random factors that best fit the data, using the *glmer()* function from the *lme4* package (ver.1.1-10) in *R* (R Core Team, 2008). Two mixed-effects logistic regression models were built, one for /i/-ɪ/ and the other for /ɛ/-æ/. Building two separate models was motivated by the prediction that the two vowel contrasts would be different in their patterns of acquisition based on their relative difficulty, as observed in previous studies (Baker et al., 2002; Ingram & Park, 2007; Tsukada et al., 2005). The resulting model for each vowel contrast included participant and item level predictors. The participant level predictor was GROUP (adults vs. children), which was centered (−0.5 and 0.5) and compared children (as the reference level, namely 0.5) with adults. The item level predictors included SPECTRUM, DURATION, and TIME. SPECTRUM and DURATION were continuous variables consisting of five spectral steps and five duration steps of the perceptual stimuli, both of which were standardized by centering and dividing by 2 standard deviations, using the *rescale()* function from the *arm* package in *R*. TIME was coded using Helmert contrasts, corresponding to Time 1 vs. Time 2 (TIME₂), the mean of the two previous time points vs. Time 3 (TIME₃), and the mean of the three previous time points vs. Time 4 (TIME₄). The motivation for using Helmert contrasts was to examine learners' improvement at each time point compared to previous time points and also to minimize collinearity as the contrasts are orthogonal. Both models included random intercepts for participants to account for participant-specific variability in responses. Random slopes for participants for SPECTRUM, DURATION, and TIME were also included for both models to account for by-participant variability in the effect of each variable on their responses (Barr, Levy, Scheeper, & Tily, 2013). Data from native English controls were not included in the statistical analysis but were included in Figures 1, 2, and 3 for comparison.

3.1.1 /i/-ɪ/

Listeners' responses for spectral and duration weighting of /i/-ɪ/ over time are shown in Figure 1 to reflect the group trend. The overall pattern of categorization responses in the upper panels illustrates that native English listeners made greater use of spectral cues than Korean learners, and that adult learners used spectrum more than child learners especially at earlier time points. As for the pattern of duration cue weighting, the lower panels in the figure show that both adult and child learners made more use of duration cues than native English listeners.

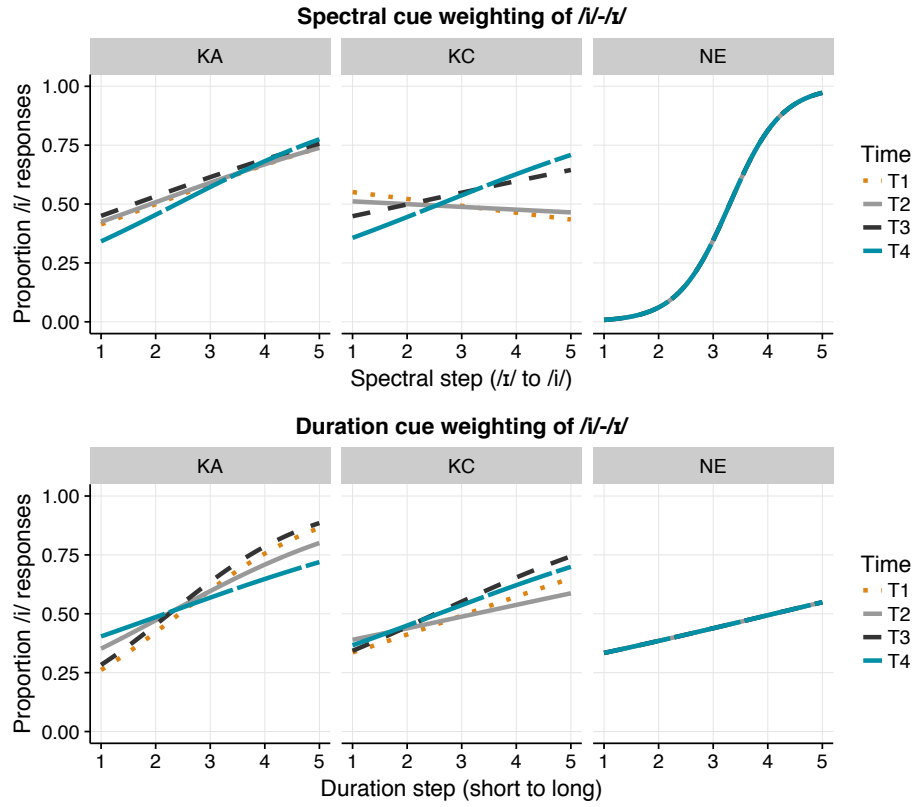


Figure 1 Proportion of /i/ responses along the vowel spectral quality continuum and the vowel duration continuum by Korean adult (KA), Korean child (KC), and native English (NE) listeners over time. Each time point is plotted by different line types and colors. The native English control group was tested only once but data are repeated at each time point for comparison.

Table 2 summarizes the estimated value for fixed-effect coefficients, along with their standard error, z statistic, and corresponding p value provided by *glmer()*.

Table 2 Summary of fixed effects for the model of the /i/-/ɪ/ contrast. Model coefficient estimates (β), standard errors, corresponding z -values, and p -values. P -values smaller than 0.05 are in bold.

Predictor	Estimate (β)	Std. Error	z	p
Intercept	0.36	0.10	3.68	< 0.001
GROUP	-0.68	0.20	-3.42	< 0.001
TIME ₂	-0.01	0.07	-0.12	0.90
TIME ₃	0.09	0.03	3.00	0.002
TIME ₄	-0.01	0.03	-0.20	0.84
SPECTRUM	1.05	0.34	3.04	0.002
DURATION	1.87	0.37	5.06	< 0.001
GROUP \times TIME ₂	0.002	0.15	-0.01	0.98
GROUP \times TIME ₃	0.02	0.06	0.35	0.72
GROUP \times TIME ₄	0.08	0.05	1.48	0.13
GROUP \times SPECTRUM	-1.20	0.69	-1.74	0.08
GROUP \times DURATION	-1.59	0.74	-2.15	0.03
TIME ₂ \times SPECTRUM	0.06	0.06	0.92	0.35
TIME ₃ \times SPECTRUM	0.19	0.04	5.24	< 0.001
TIME ₄ \times SPECTRUM	0.23	0.03	8.43	< 0.001
TIME ₂ \times DURATION	-0.34	0.06	-5.32	< 0.001
TIME ₃ \times DURATION	0.19	0.04	4.88	< 0.001

TIME ₄ × DURATION	−0.13	0.03	−4.70	< 0.001
GROUP × TIME ₂ × SPECTRUM	0.23	0.13	1.76	0.07
GROUP × TIME ₃ × SPECTRUM	0.34	0.08	4.48	< 0.001
GROUP × TIME ₄ × SPECTRUM	0.22	0.06	3.86	< 0.001
GROUP × TIME ₂ × DURATION	0.27	0.13	2.02	0.04
GROUP × TIME ₃ × DURATION	−0.001	0.08	−0.01	0.99
GROUP × TIME ₄ × DURATION	0.32	0.06	5.76	< 0.001

Overall, listeners responded /i/ 52% of the time, indicating a slight tendency towards /i/ responses ($\beta = 0.36$, $z = 3.68$, $p < 0.001$). The model found significant main effects of GROUP, TIME₃, SPECTRUM, and DURATION, indicating that /i/ vowel responses were overall higher for adult than child learners, and at Time 3 than at previous time points. More importantly, Korean listeners gave more /i/ vowel responses as the vowel spectral ($\beta = 1.05$, $z = 3.04$, $p = 0.002$) and duration steps shift to the /i/ vowel ($\beta = 1.87$, $z = 5.06$, $p < 0.001$), indicating that they are sensitive to spectral and duration changes.

Significant two- and three-way interactions between each of the cues (SPECTRUM and DURATION), GROUP and TIME indicate that use of the cues changed over time and did so differently for the two age groups. The effect of SPECTRUM increased at Time 3 ($\beta = 0.19$, $z = 5.24$, $p < 0.001$) and Time 4 ($\beta = 0.23$, $z = 8.43$, $p < 0.001$) and these increases were larger for the child learners (GROUP × TIME₃ × SPECTRUM, $\beta = 0.34$, $z = 4.48$, $p < 0.001$; GROUP × TIME₄ × SPECTRUM, $\beta = 0.22$, $z = 3.86$, $p < 0.001$). Conversely, the effect of DURATION decreased from Time 1 to Time 2 ($\beta = -0.34$, $z = -5.32$, $p < 0.001$), increased at Time 3 ($\beta = 0.19$, $z = 4.88$, $p < 0.001$), and decreased again at Time 4 ($\beta = -0.13$, $z = -4.70$, $p < 0.001$). These decreases were also larger for the child learners (GROUP × TIME₂ × DURATION, $\beta = 0.27$, $z = 2.02$, $p = 0.04$; GROUP × TIME₄ × DURATION, $\beta = 0.32$, $z = 5.76$, $p < 0.001$). Taken together, these results indicate that learners decreased their reliance on duration and increased their reliance on spectral cues over time and that these changes were greater for the children.

3.1.2 /ε/-/æ/

Figure 2 shows the group results representing listeners' responses varying by spectral and duration steps for /ε/-/æ/ over time. As with /i/-/ɪ/ above, the overall pattern of categorization responses in the upper panels demonstrates that native English listeners used spectral cues much more than Korean learners in classifying stimuli as /ε/ or /æ/. For duration, the lower panels of Figure 2 show that although both adult and child learners gave more /ε/ vowel responses initially for longer vowels (the opposite pattern from native English listeners), their use of duration cues at Time 4 was similar to native English listeners.

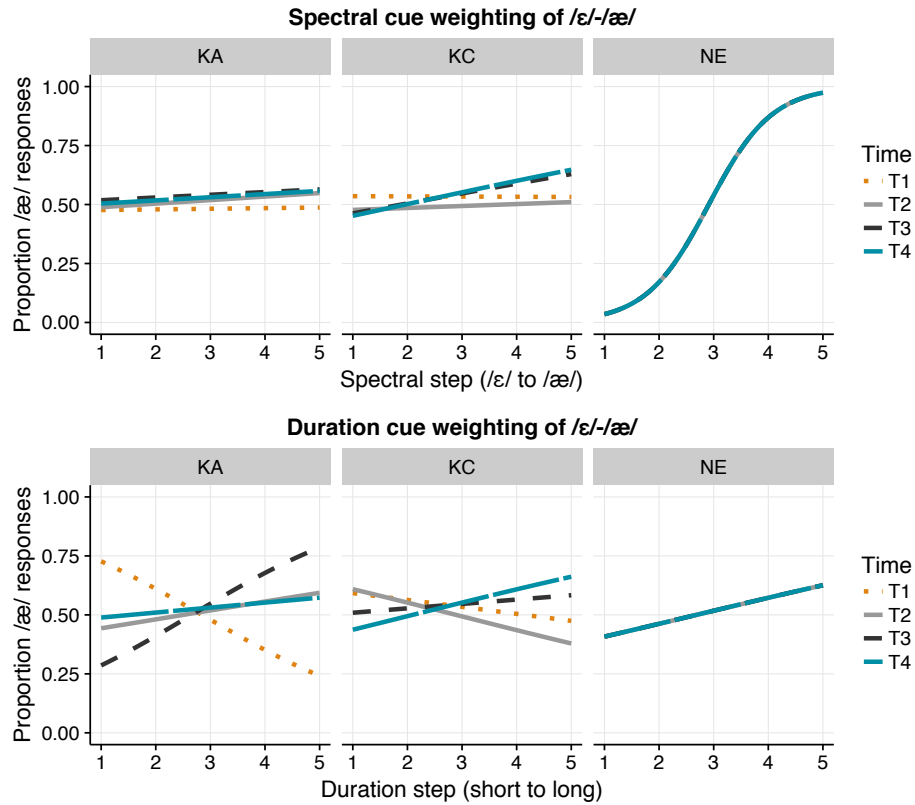


Figure 2 Proportion of /æ/ responses along the vowel spectral quality continuum and the vowel duration continuum by Korean adult (KA), Korean child (KC), and native English (NE) listeners over time. Each time point is plotted by different line types and colors. Data from the native English control group are repeated each time point for comparison.

The results from the mixed-effects logistic regression are summarized in Table 3 as the estimated value for fixed-effect coefficients, along with their standard error, z statistic, and corresponding p value.

Table 3 Summary of fixed effects for the /ɛ/-/æ/ contrast model. Model coefficient estimates (β), standard errors, corresponding z -values, and p -values. P -values smaller than 0.05 are bolded.

Predictor	Estimate (β)	Std. Error	z	p
Intercept	0.11	0.05	2.31	0.02
GROUP	0.05	0.10	0.49	0.62
TIME ₂	-0.004	0.04	-0.08	0.93
TIME ₃	0.07	0.03	2.11	0.03
TIME ₄	0.03	0.03	1.08	0.28
SPECTRUM	0.25	0.12	2.04	0.04
DURATION	0.11	0.28	0.38	0.70
GROUP \times TIME ₂	-0.19	0.09	-2.19	0.02
GROUP \times TIME ₃	-0.04	0.07	-0.58	0.56
GROUP \times TIME ₄	0.02	0.05	0.30	0.76
GROUP \times SPECTRUM	0.17	0.25	0.68	0.49
GROUP \times DURATION	-0.34	0.57	-0.60	0.55
TIME ₂ \times SPECTRUM	0.07	0.05	1.31	0.19
TIME ₃ \times SPECTRUM	0.11	0.03	3.30	< 0.001
TIME ₄ \times SPECTRUM	0.07	0.02	2.81	0.004
TIME ₂ \times DURATION	0.46	0.06	7.96	< 0.001
TIME ₃ \times DURATION	0.57	0.03	16.70	< 0.001
TIME ₄ \times DURATION	0.18	0.02	7.32	< 0.001

GROUP \times TIME ₂ \times SPECTRUM	-0.04	0.11	-0.38	0.70
GROUP \times TIME ₃ \times SPECTRUM	0.14	0.06	2.22	0.02
GROUP \times TIME ₄ \times SPECTRUM	0.11	0.04	2.29	0.02
GROUP \times TIME ₂ \times DURATION	-1.43	0.12	-12.31	< 0.001
GROUP \times TIME ₃ \times DURATION	-0.62	0.07	-8.96	< 0.001
GROUP \times TIME ₄ \times DURATION	0.22	0.05	4.65	< 0.001

Overall, listeners responded /æ/ 55% of the time, indicating a slight tendency towards /æ/ responses ($\beta = 0.11$, $z = 2.31$, $p = 0.02$). The model found significant main effects of TIME₃ and SPECTRUM. This indicates that Korean listeners made more /æ/ responses at Time 3 than at the previous time points ($\beta = 0.07$, $z = 2.21$, $p = 0.03$) and that Korean listeners overall were influenced by the spectral changes ($\beta = 0.25$, $z = 2.04$, $p = 0.04$). The lack of an overall duration effect likely reflects the change in the direction of the duration effect between the first two time points observed in Figure 2.

The model also found significant two- and three-way interactions between each of the cues (SPECTRUM and DURATION), GROUP, and TIME, indicating that use of the cues changed over time and did so differently for adult and child learners. The effect of SPECTRUM increased at Time 3 ($\beta = 0.11$, $z = 3.30$, $p < 0.001$) and Time 4 ($\beta = 0.07$, $z = 2.81$, $p = 0.004$) and these increases were larger for the child learners (GROUP \times TIME₃ \times SPECTRUM, $\beta = 0.14$, $z = 2.22$, $p = 0.02$; GROUP \times TIME₄ \times SPECTRUM, $\beta = 0.11$, $z = 2.29$, $p = 0.02$). On the other hand, the effect of DURATION decreased from Time 1 to Time 2 ($\beta = 0.46$, $z = 7.96$, $p < 0.001$), increased at Time 3 ($\beta = 0.19$, $z = 4.88$, $p < 0.001$), and decreased again at Time 4 ($\beta = -0.13$, $z = -4.70$, $p < 0.001$). The effect of DURATION increased more for the adult learners at Time 2 and Time 3 (GROUP \times TIME₂ \times DURATION, $\beta = -1.43$, $z = -12.31$, $p < 0.001$; GROUP \times TIME₃ \times DURATION, $\beta = -0.62$, $z = -8.96$, $p < 0.001$), but it increased more for the child learners at Time 4 (GROUP \times TIME₄ \times DURATION, $\beta = 2.22$, $z = 4.65$, $p < 0.001$). Overall, as can be seen in Figure 2, these results indicate that learners mostly changed their cue weights for duration over time. Both age groups changed their cue weights for duration over time but this effect was found earlier for adult than child learners. For the spectral cues, child learners made small but significant changes at later time points.

When comparing the two vowel contrasts, the overall group results indicate different patterns of development between /i/-/ɪ/ and /ɛ/-/æ/ over the course of a year. Notably, learners increased their reliance on spectral cues for the /i/-/ɪ/ contrast whereas their improvement was mostly limited to the duration dimension for the /ɛ/-/æ/ contrast. For /i/-/ɪ/, both adult and child learners changed their cue weights for both dimensions over time, and the improvement was especially noticeable in the spectral dimension for the child learners. For /ɛ/-/æ/, however, they showed no or minimal improvement in the spectral dimension although both learner groups made improvements in the duration dimension.

3.2 Individual results

The results presented in the preceding section average across all participants. Although examining group trends can illuminate how adult and child learners differ in their acquisition patterns, the longitudinal approach taken in the present study allows for a detailed investigation into individuals' developmental trajectories for L2 vowel acquisition. As will be shown below, a close inspection of the data from individual participants reveals that learners make different use of acoustic cues in classifying the vowel contrasts: some learners initially make more use of duration cues whereas other learners are more attuned to spectral cues.

3.2.1 Individual cue weights

Logistic regression coefficients were used to quantify how listeners use the available acoustic cues (Morrison, 2005, 2007; Morrison & Kondaurova, 2009). A series of logistic regression models were fitted to each listener's proportion of /i/ and /æ/ responses. Individual participants' perceptual weights for spectral and duration cues were calculated based on spectral and duration

coefficients fitted to each listener's response data. The coefficients from the individual models were used as measures of the perceptual weight of the respective cues. We used this method, rather than using random slopes by participant from the mixed effects model, because we did not expect individuals' deviations from the mean coefficient to be normally distributed as is assumed for random slopes in a mixed-effects model. Figure 3 shows a scatterplot of the coefficient values. Participants' spectral and duration reliance were shown as the spectrally-tuned logistic regression coefficient (β_{spec}) and duration-tuned logistic regression coefficient (β_{dur}) values, respectively. Negative cue weights indicate that listeners were influenced by the cue but in the opposite direction from native listeners.

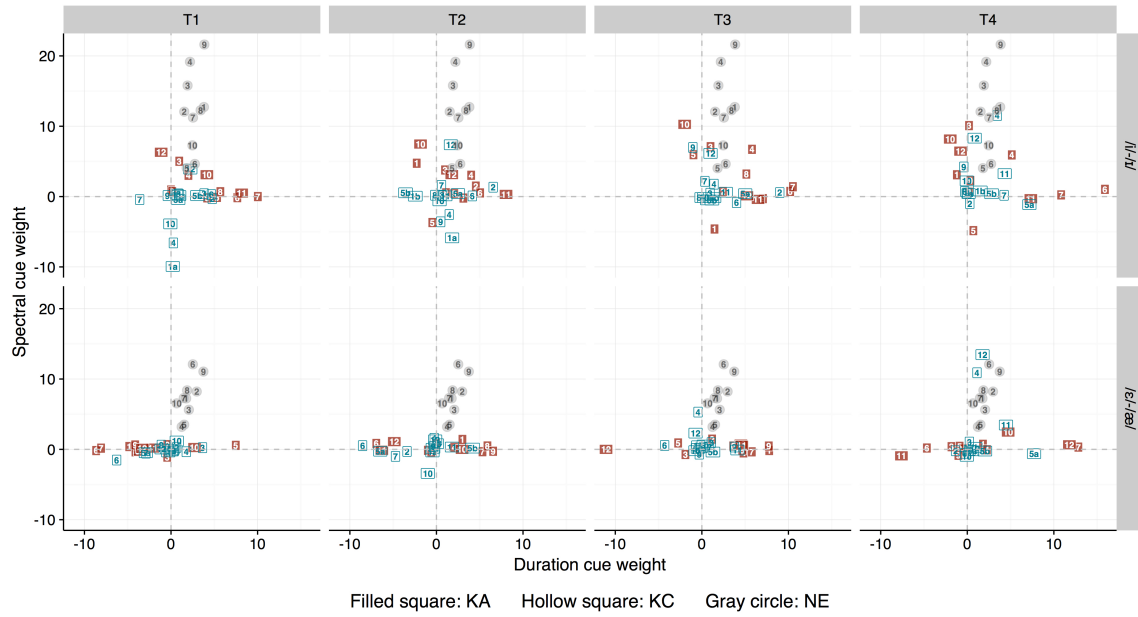


Figure 3 Scatter plots of the reliance on spectral and duration cues by groups across time points, as indicated by the coefficients from the logistic regression analysis fitted to each listener's response data. KA = Korean Adults, KC = Korean Children, NE = Native English. Data from the native English control group are repeated at each time point for comparison.

Figure 3 illustrates native English listeners' high β_{spec} values and low β_{dur} values, indicating that they relied primarily on spectral information in categorizing the target vowels for both vowel contrasts, as expected. Korean learners, however, differed from the native English control group in their use of spectral and duration cues. Individual learners' cue weights for each vowel contrast are provided in the appendix (Tables B1, B2). Notably, most learners in Figure 3 are located either on the x-axis or on the y-axis, not in the middle of one of the quadrants. That is, although some Korean listeners used both spectral and duration information in classifying the target vowel contrasts, most only used one cue or the other rather than a combination of both cues.

For the /i/-/ɪ/ contrast, the Korean learners were initially widely scattered with both negative and positive β_{spec} and β_{dur} values. Learners shifted their β_{spec} and β_{dur} values towards more positive values over time. Positive β_{spec} values were more associated with adult learners and positive β_{dur} values were more associated with child learners at early time points (Times 1 and 2). More adults and children had positive β_{spec} values over time. In other words, more adult learners than child learners initially used spectral cues, but over time more child learners increased their spectral reliance. Although there were still quite a few learners who consistently used duration cues at Time 4, all learners used duration cues in the native-like direction (i.e., longer vowel duration for /i/) from Time 3.

For the /ε/-/æ/ contrast, all learners' coefficients were initially distributed along the duration dimension. β_{dur} values shifted towards more positive β_{dur} values over time. From Time 3, a few child learners started utilizing spectral cues and they enhanced their use of spectral cues at Time 4 with much bigger β_{spec} values. Notably, only child learners showed this pattern of development on the spectral dimension and the majority of learners continued to use duration cues over time.

In sum, the analyses of the individual participants' perceptual weights revealed considerable variation in terms of the use of spectral and duration cues and that most learners tended to use only one acoustic dimension. The results show that most Korean learners were attuned to duration cues for both vowel contrasts while only some Korean learners made use of spectral cues especially for the /i/-/ɪ/ contrast. It is particularly remarkable that the use of spectral cues for the /ε/-/æ/ contrast was only achieved by a few child learners. These patterns are in line with the group analyses in Section 3.1.2, but show that changes to the group data over time were largely driven by a handful of individuals.

3.2.2 Grouping by initial state

To complement the quantitative analysis of raw cue weights above, we also performed a grouping analysis that divided Korean learners into different groups based on their cue weights at Time 1. More specifically, we grouped participants on the basis of whether the logistic regression coefficients (cue weights) for each dimension were significantly different from zero (see Tables B1 and B2 in the appendix for individual weights and statistical significance). We also used a hierarchical cluster analysis to determine the number of groups. Although hierarchical cluster analysis is an exploratory procedure with some degree of subjectivity, it helps identify groups in a relatively controlled way (Morrison, 2008). Grouping this way allowed us to further examine the trajectories of individuals who were in the same group at Time 1.

Our grouping process yielded six initial groups for the /i/-/ɪ/ contrast (Figure 4) and three groups for the /ε/-/æ/ contrast (Figure 5). The characteristics of different groups are described as follows: an initial “-Dur” group that used duration to distinguish the vowel contrasts but with negative β_{dur} values (i.e., learners made fewer /i/ or /æ/ vowel responses as vowel duration increased), an initial “Dur” group that used duration to distinguish the vowel contrasts with positive β_{dur} values, an initial “Spec+Dur” group that used both spectrum and duration to distinguish the vowel contrasts, an initial “-Spec” group that used spectrum to distinguish the vowel contrasts but with negative β_{spec} values (i.e., learners made fewer /i/ or /æ/ vowel responses as spectral steps shifted to /i/ or /æ/), an initial “Spec” group that primarily used spectrum to distinguish the vowel contrasts, and finally an initial “Random” group that randomly chose stimuli with no distinction between the two vowels yielding cue weights close to zero in both dimensions. With the exception of the “-Spec” and “-Dur”, these different initial cue weighting strategies correspond to the different stages assumed in previous studies (Escudero, 2000; Morrison, 2008), which we will return to in the next section.

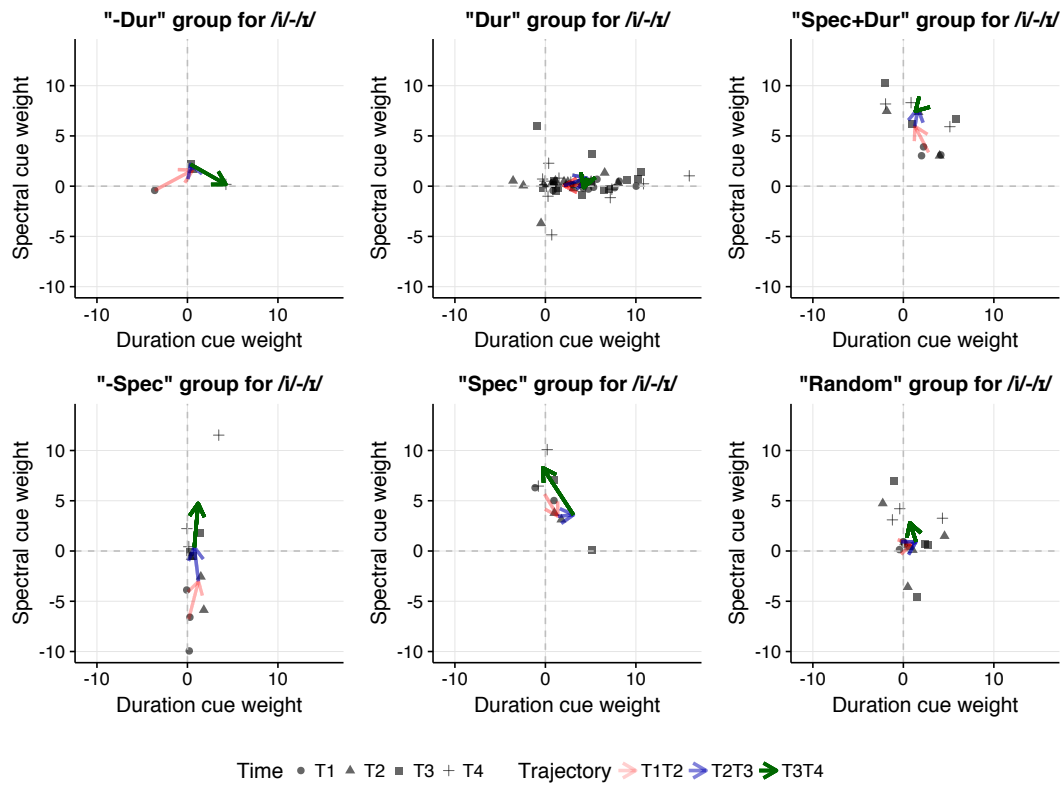


Figure 4 Scatter plots of the reliance on spectral and duration cues for /i/-/ɪ/ by different groups based on their initial cue weighting patterns, as indicated by the coefficients from the logistic regression analysis fitted to each listener's response data, and the hierarchical cluster analysis. Developmental trajectories are indicated by arrows where the tail of T1T2 is the mean coefficients at Time 1, the head of T1T2 and the tail of T2T3 are the mean coefficients at Time 2 (Trajectory T1T2 in light red), the head of T2T3 and the tail of T3T4 are the mean coefficients at Time 3 (Trajectory T2T3 in blue), and the head of T3T4 is the mean of the coefficients at Time 4 (Trajectory T3T4 in dark green).

Figure 4 illustrates the six groupings for the /i/-/ɪ/ contrast based on the initial cue weighting strategies of the learners. The directions of the developmental trajectories over time, indicated by arrows, show that the groups that initially relied exclusively on duration-based cues—the initial “-Dur” (1 child) and “Dur” groups (6 adults & 7 children)—progressed along the duration dimension over time. In the same way, the groups that relied most on spectral-based cues initially—the initial “-Spec” (3 children) and “Spec” groups (2 adults)—progressed mostly along the spectral dimension over time. A similar pattern also held for the initial “Spec+Dur” group (2 adults & 1 child). On average, the initial “Random” group (2 adults & 2 children) did not move in any particular direction or in any systematic way.

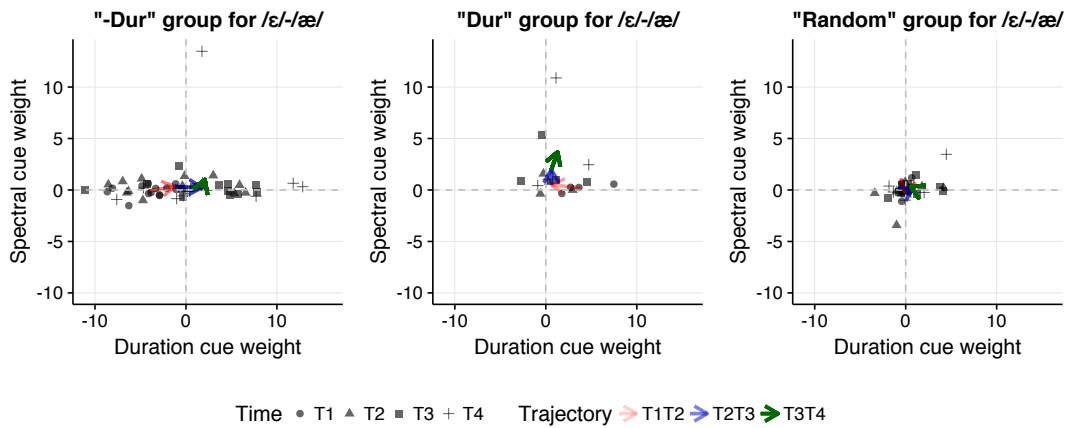


Figure 5 Scatter plots of the reliance on spectral and duration cues for /ɛ/-/æ/ by different groups based on their initial cue weighting patterns, as indicated by the coefficients from the logistic regression analysis fitted to each listener’s response data, and the hierarchical cluster analysis. Developmental trajectories are indicated by arrows of which the tail of T1T2 is the mean coefficients at Time 1, the head of T1T2 and the tail of T2T3 are the mean coefficients at Time 2 (Trajectory T1T2 in light red), the head of T2T3 and the tail of T3T4 are the mean coefficients at Time 3 (Trajectory T2T3 in blue), and the head of T3T4 is the mean of the coefficients at Time 4 (Trajectory T3T4 in dark green).

Figure 5 illustrates the three groupings for the /ɛ/-/æ/ contrast based on the initial cue weighting strategies of the learners which only included “-Dur”, “Dur”, and “Random”. A large number of Korean learners were in the initial “-Dur” group (8 adults & 5 children). As a whole, this group moved towards more native-like perception for /ɛ/-/æ/ with positive β_{dur} values. Interestingly, some learners in the initial “Dur” group (2 adults & 2 children) started to use spectral cues at later time points, moving the group trajectory towards spectral reliance. As in the /i/-/ɪ/ contrast, the initial “Random” group (2 adults & 7 children) for the /ɛ/-/æ/ contrast did not move in any particular direction or in any systematic way. Thus, the patterns we observe for the /ɛ/-/æ/ contrast are largely consistent with /i/-/ɪ/. The most noticeable difference between the two contrasts is that initially no Korean learners used spectrum to distinguish the /ɛ/-/æ/ contrast while some learners showed spectrum-based patterns for the /i/-/ɪ/ distinction. This means that the same learners who were initially sensitive to spectrum for the high vowels were not for the low vowels. We also see that in the early stages learners’ cue weighting for the /ɛ/-/æ/ contrast diverges more from that of native English listeners than the weighting for the /i/-/ɪ/ contrast, suggesting that the acquisition of the /ɛ/-/æ/ contrast not only lags behind when compared to the /i/-/ɪ/ contrast, but it may start at a different point.

In summary, by grouping learners according to their initial states, we found six initial states for the /i/-/ɪ/ contrast including reliance on one cue or the other, reliance on both cues, and some who seemed to rely on neither (i.e., “Random”). For the /ɛ/-/æ/ contrast, we found only three initial states, as no learners initially relied on the spectral cue. The group trajectories for these different initial states over time were also different, with learners tending to keep to one dimension. However, as is clear from Figures 4 and 5, even these groupings by initial state include quite a bit of individual variability in trajectories over time, which we turn to next.

3.2.3 Individual trajectories in perceptual cue weighting

Our final analysis included plotting individual trajectories over time. Figure 6 displays individual learners’ developmental trajectories in perceptual cue weighting across the vowel contrasts over time. We also categorized learners at each time point into the same groups as in the previous section. Individual learners’ cue weighting patterns for each vowel contrast at each time point are provided in the appendix (Table B3). The developmental paths for individual learners were examined to determine whether individual trajectories are consistent with the developmental stage hypothesis in Escudero (2000), as discussed in the introduction.

Furthermore, the individual trajectories were analyzed to more carefully probe the “Random”, “–Dur”, and “–Spec” patterns.

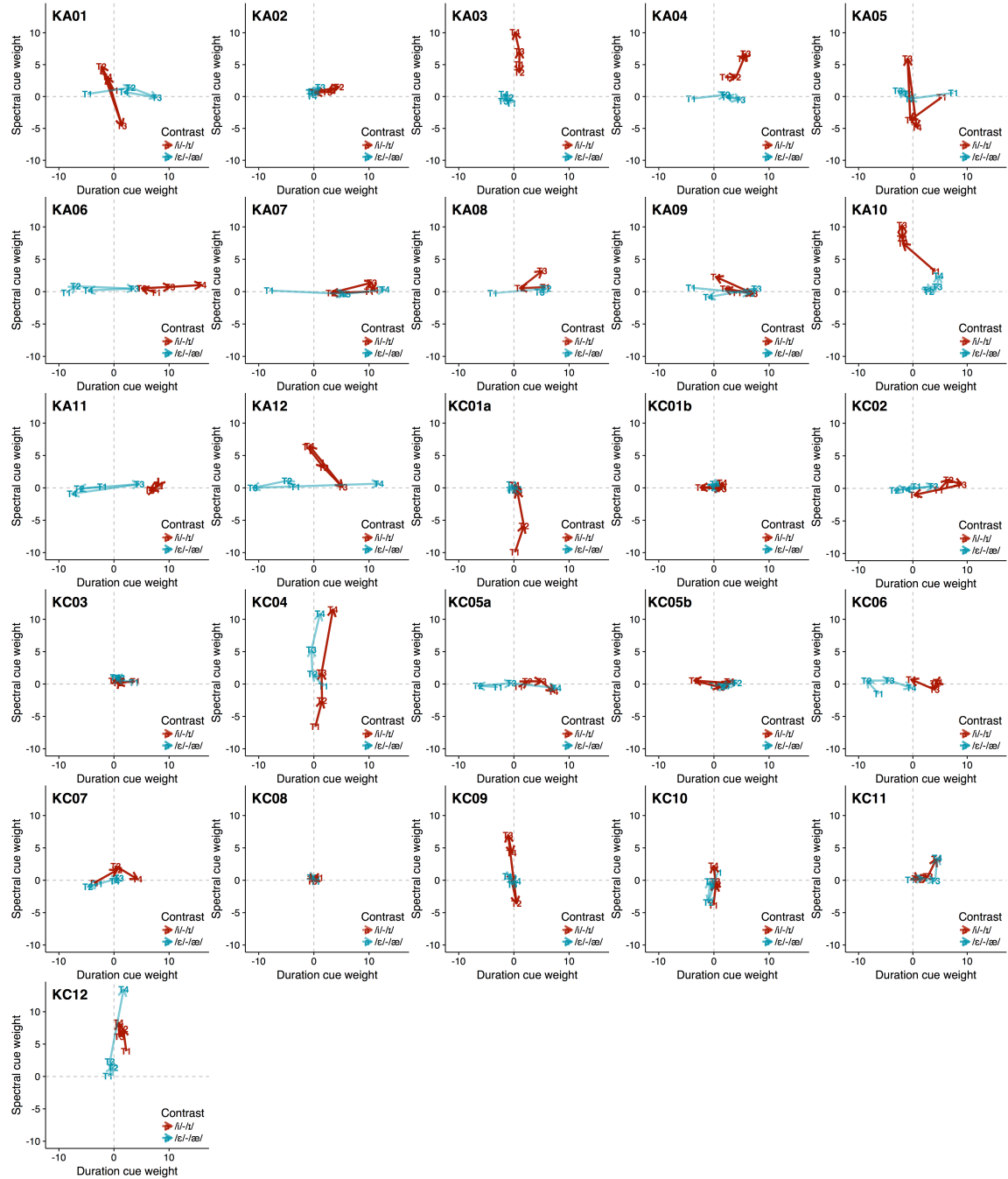


Figure 6 Developmental trajectories across time points for individual learners. Each time point is indicated by the coefficient from the logistic regression analysis fitted to each learner’s response data. The vowel contrasts are displayed in different colors (/i/-/ɪ/ in dark red and /ε/-/æ/ in light blue), and the developmental trajectories of individuals are illustrated by arrows connecting Time 1 with Time 2, Time 2 with Time 3, and Time 3 with Time 4.

As discussed in the introduction, the developmental stage hypothesis states that in the initial stage learners are not able to identify tokens of /i/ versus /ɪ/, thereby showing a no-contrast pattern (“Random” in our groupings). At the next stage, the contrast is distinguished by duration information exclusively (“Dur”). Then, learners use both duration and spectral

information but still give priority to duration cues (“Spec+Dur”). At the final stage, learners show English-like use of both spectral and duration cues with spectral cues receiving primary weighting (“Spec”) as for native listeners (Escudero, 2000). While hypotheses were not provided by Escudero for other vowel contrasts, we might expect /ε/-/æ/ to follow a similar pattern, ending with more reliance on spectral cues.

Inspection of the trajectories in Figure 6 and the appendix reveals that for the /i/-/ɪ/ contrast, the overall direction of change over time is consistent with Escudero’s (2000) stages of development (i.e., duration to spectrum) for 8 learners (KA05, KA08, KA09, KA12, KC01b, KC02, KC11, KC12). However, there were 7 learners who did not follow this pattern. These learners did not go through a duration stage but rather showed spectral reliance from earlier time points (KA01, KA03, KA10, KC04, KC09, KC10, KC12). We also found that for learners who reached the “Spec” stage, their development was immediately preceded by “Spec+Dur” (KA03, KA10, KC12), “-Spec” (KA05, KC09), “Dur” (KA09), “-Dur” (KC07), or “Random” (KC10) for /i/-/ɪ/. Only “Spec+Dur” was hypothesized to come directly before “Spec”. For the /ε/-/æ/ contrast, most of the learners who reached a spectrum-related stage went through an earlier duration stage (KA10, KA11, KC03, KC03, KC11).

The “Random” category requires special attention. If this represents the earliest stage of development in which learners are not sensitive to either cue, as hypothesized by Escudero (2000), we might expect that learners who start in this category take more time to arrive at the native-like cue weighting pattern. This was the case for some learners who remained stagnant in terms of the developmental trajectories observed over time (KA02, KC01a, KC01b, KC08 for /ε/-/æ/): these learners were not sensitive to either spectral or duration cues and their developmental trajectories did not change much over time. However, other learners in the initial “Random” group were somewhat unpredictable at later time points. That is, some went from “Random” to “-Dur” (KA03, KC02 for /ε/-/æ/), and others from “Random” to “-Spec” (KC09 for /i/-/ɪ/, KC10 for /ε/-/æ/). We even observed at some points that some learners went from “Dur” to “Random” or from “-Spec” to “Random”. For example, for the /i/-/ɪ/ contrast, some learners moved from “Dur” to “Random” (KA02, KC03, KC06, KC08), and others moved from “-Spec” to “Random” (KC01a) then to “Spec” (KC10). For the /ε/-/æ/ contrast, “Dur” to “Random” (KA05, KC07), “-Dur” to “Random” (KC05a, KC06, KC09), and even subtle use of “Spec+Dur” to “Random” (KA02) patterns were found along the developmental trajectories. This suggests that although for some learners the “Random” behavior might mean inability to reliably distinguish the contrast, others may in fact be sensitive to certain cues but be unsure how to map them onto the categories, and therefore show the “Random” pattern in a transition period of development. In this process, they may have tried different perceptual strategies and fluctuated between different cues but failed to systematically adopt relevant perceptual strategies for the contrast. Thus, when learners are categorized as “Random” in cross-sectional studies, this might not always indicate that the learners in this group are insensitive to certain cues, and these patterns should be treated with some caution.

The most surprising patterns, which were not addressed in Escudero (2000) and Morrison (2008), are the initial “-Dur” and “-Spec” groups which interpreted the cues in the opposite direction from native listeners. In fact, negative values for spectral reliance were found among a few learners in Escudero (2000), but they were interpreted as being due to chance. Flege et al. (1997) also reported reversals which in some cases they thought were due to orthography. One possible reason for these patterns is that the learners initially mislabeled the words. That is, the learners might have thought that the concept ‘bet’ sounded like /bæt/ and the concept ‘bat’ sounded like /bet/. Most of the learners who initially used “-Dur” in fact shifted to “Dur” at later time points (KC01b, KC05b for /i/-/ɪ/ and KA04, KA06, KA07, KA08, KA09, KA11, KA12, KC02, KC05a, KC07 for /ε/-/æ/), which is in line with this initial mislabeling account. The “-Spec” stage preceded “Spec” (KA05, KC09, KC10), “Spec+Dur” (KA01, KC04), or “Random” (KC01a, KC09) for /i/-/ɪ/, and “-Dur” (KC10), “Dur” (KA11), or “Random” (KC10) for /ε/-/æ/ making this finding harder to interpret. Overall, these results suggest that individual trajectories are not uniform, and that the direction and the rate of development can also be modulated by different vowel contrasts.

4.0 Discussion

The present longitudinal study has examined the acquisition of novel vowel contrasts by Korean learners of English in the early stages of acquisition in an L2-speaking environment. To our knowledge, this is the first longitudinal investigation of developmental trajectories in cue weighting of L2 vowel contrasts. Our results confirm earlier findings (Casillas, 2015; Escudero, 2000, 2005; Escudero et al., 2009; Flege et al., 1997; Kondaurova & Francis, 2008, 2010; Morrison, 2008) that L2 learners can learn to prioritize more important acoustic dimensions relevant to English vowel contrasts over time. As hypothesized (Escudero, 2000, 2005; Morrison, 2008), the learners were initially more influenced by duration cues, but their reliance on spectral cues, which are relevant to the target phonetic categories, gradually increased over time. The results also reveal that the learners' perceptual patterns were modulated by vowel contrast as in earlier studies (Baker et al., 2002; Bohn & Flege, 1990; Ingram & Park, 1997; Tsukada et al., 2005). We found that the relevant acoustic dimensions were learned earlier for the /i/-/ɪ/ contrast. However, individuals differed in their cue weighting strategies such that some learners relied primarily on duration cues, while others were more influenced by spectral cues, even at Time 1. The rate of learning to use the relevant acoustic cues (i.e., spectral cues) also differed across learners.

4.1 Individual differences in L2 speech perception

The results of this study indicate that individual L2 learners differ widely in their developmental trajectories of perceptual cue weighting over time. The longitudinal results revealed that seemingly random individual variability in perceptual cue weighting can be understood as a snapshot from a developmental sequence along which individual learners are situated. That is, learners' cue weighting strategies can be predicted in terms of which acoustic cues they use and how they progress as they become more native-like listeners in English. This developmental pattern is in line with previous work (e.g., Mayr & Escudero, 2010), which focused on the development of L2 vowel perception. However, the present study differs from previous work in that the systematicity of the perception patterns by individual learners was explained by the detailed longitudinal trajectories rather than the degree of variation in performance in a cross-sectional design.

One potential account of the different perceptual strategies shown by individual learners might be their auditory abilities (Lengeris, 2009; Wong & Perrachione, 2007). In his study on individual differences in the learning of English vowels by Greek speakers, Lengeris (2009) showed that success in L2 vowel learning was related to individuals' auditory abilities evaluated in non-speech psychoacoustic tasks. Wong and Perrachione (2007) examined the learning of non-native pitch patterns for word identification by L2 learners and found that successful learning of pitch patterns in lexical identification was predicted by individual learners' auditory abilities to perceive different pitch patterns. The findings of these studies suggest that individual differences in L2 speech perception might be related to general auditory acuity.

An alternative explanation is that individual differences in perceptual assimilation of L2 sounds to L1 categories are responsible for L2 perceptual development (e.g., as argued by Mayr & Escudero, 2010). In other words, some individuals might assimilate two L2 sounds to a single L1 category to a greater degree than others. Hattori and Iverson (2009), however, argued against this view. They examined whether individual Japanese learners' ability to perceive and produce English /ɪ/ and /I/ is related to the individual's degree of assimilation of /ɪ/ and /I/ into the Japanese /ɪ/ category. That is, they aimed to determine whether L2 learning difficulties stem from how L2 sound categories are assimilated into one L1 category. They found that Japanese learners' discrimination of English /ɪ/-/I/ was not readily predicted by the degree of assimilation of English /ɪ/-/I/ into Japanese /ɪ/, but rather by how much individual learners used a particular acoustic dimension (F3) to identify the English /ɪ/-/I/ category. Thus, they argued that individual differences in cue weighting are more tightly linked to proficiency than to assimilation patterns. Flege et al. (1997) also found a close link between cue weights and production proficiency.

Regarding individual L2 learners' use of acoustic cues in phonetic categorization, Schertz et al. (2016) showed that variability in cue weighting strategies also contributes to L2 learners' adaptation of L2 phonetic categories. Kong and Edwards (2015) provided additional evidence that L2 learners' cue weighting strategies are not uniform and suggested that individual cue weighting strategies in the L1 along with L2 proficiency are contributing factors to individual differences in L2 perceptual cue weighting. The findings of the present study corroborate these previous studies suggesting that examining individual learners' cue weighting patterns sheds important light on individual differences in L2 speech perception. The present findings further suggest that differences in learners' initial cue weighting strategies, which might indicate baseline differences in their abilities to use acoustic-phonetic information in the speech signal, are related to the direction and the rate of development in L2 speech perception over time.

4.2 Differential acquisition patterns between the vowel contrasts

The present study is consistent with previous findings that L2 learners' perceptual patterns can be modulated by different vowel contrasts (Baker et al., 2002; Bohn & Flege, 1990; Flege et al., 1997; Ingram & Park, 1997; Tsukada et al., 2005). In the present study, however, we confirmed the relative difficulty between the two vowel contrasts by examining the developmental trajectories of each contrast longitudinally. This study found that L2 learners showed different acquisition patterns of cue weighting between the two vowel contrasts and that these patterns can shed light on the relative difficulty between the two contrasts. That is, the Korean learners used both spectral and duration cues to distinguish the /i/-/ɪ/ contrast, but primarily used duration for the /ɛ/-/æ/ contrast. The Korean learners were able to use spectral differences from Time 1 to distinguish /i/-/ɪ/ but most of the learners failed to use spectral differences for the /ɛ/-/æ/ contrast until Time 3. Additionally, the individual trajectories showed that acquisition of the /ɛ/-/æ/ contrast lags behind compared to the /i/-/ɪ/ contrast, suggesting that /ɛ/-/æ/ is relatively more difficult. This is in line with previous research which similarly found that the high vowel contrast was easier to acquire for Korean learners of English (Baker et al., 2002; Ingram & Park, 1997; Tsukada et al., 2005) as well as for learners from other L1 backgrounds (Baptista, 2006; Bohn & Flege, 1990; Jia, Strange, Wu, Collado, & Guan, 2006).

Previous L2 speech learning models make predictions for perceptual difficulties in the discrimination of non-native contrasts. The PAM/PAM-L2, in particular, predicts levels of difficulty in discrimination of L2 contrasts based on the patterns of perceptual assimilation of L2 segments to native language categories (Best, 1995; Best & Tyler, 2007). For example, when two L2 sounds are assimilated to two different L1 speech sounds (two-category assimilation), these two sounds should be well discriminated. When two L2 sounds are assimilated to one L1 category but one L2 sound is a better exemplar of the L1 category than the other, discrimination is predicted to be poorer than a case of two-category assimilation (category goodness assimilation). Discrimination will be problematic when both L2 sounds are assimilated to the same L1 category equally well or poorly (single-category assimilation). The perceptual mapping of L1 and L2 sounds is also used in Flege's (1995) Speech Learning Model to predict successful learning of L2 sounds (e.g., perception accuracy). The Second Language Perception Model (Escudero, 2005; Escudero & Vasiliev, 2011; van Leussen & Escudero, 2015) similarly predicts that initial assimilation patterns should determine how difficult a contrast is to learn. Some studies on perceptual assimilation with inexperienced Korean listeners have found that the English vowel contrasts /i/-/ɪ/ and /ɛ/-/æ/ are assimilated to a single Korean vowel, /i/ and /ɛ/, respectively (Ingram & Park, 1997; Son, 2008; Yun, 2014). On closer inspection, however, these studies have shown marginally different trends in the assimilation patterns between the two vowel contrasts in which English /ɛ/ and /æ/ are assimilated to Korean /ɛ/ to comparable extents but English /i/ is somewhat more assimilated to Korean /i/ than English /ɪ/ is. Similarly, other studies have reported that native Korean listeners show different assimilation patterns between English /i/-/ɪ/ and /ɛ/-/æ/ (Baker et al., 2002; Tsukada et al., 2005). In particular, Baker et al. (2002) showed that native adult Korean listeners chose Korean /i/ for English /i/ 92% of the time and /ɪ/ 68% of the time whereas they chose Korean /ɛ/ for English /ɛ/ 57% of the time

and /æ/ 67% of the time. In this case, English /i/ is better matched to Korean /i/ than English /ɪ/ is whereas both English /ɛ/ and /æ/ correspond to Korean /ɛ/ relatively poorly. This indicates that the /i/-/ɪ/ contrast is likely a case of category goodness assimilation whereas the /ɛ/-/æ/ contrast is a case of single-category assimilation.³ Thus, these findings suggest one possible explanation for why Korean learners are likely to have more difficulty with the acquisition of English /ɛ/-/æ/ than with /i/-/ɪ/.

Another account for the relative difficulty between /i/-/ɪ/ and /ɛ/-/æ/ derives from the acoustic distinctiveness of the cues to the two vowels in each contrast in English. First, the spectral difference between /i/ and /ɪ/ may be larger than the spectral difference between /ɛ/ and /æ/. While the /i/-/ɪ/ and /ɛ/-/æ/ contrasts involve a fairly similar difference in height, an F1 difference of approximately 150 Hz, the /i/-/ɪ/ contrast involves a much larger difference in frontness, with an F2 difference of about 500 Hz compared to approximately 160 Hz for /ɛ/-/æ/ in Canadian English (Boberg, 2010). Furthermore, the distribution of F1 and F2 values overlap more for the /ɛ/-/æ/ vowels (Fridland, Kendall, & Farrington, 2014; Hillenbrand et al., 2000). Hillenbrand et al. (2000) reported that when they modified vowel duration in vowel classification tests, the recognition of /ɛ/-/æ/ was strongly influenced by vowel duration modification whereas modifying vowel duration had a minimal effect on the recognition of /i/-/ɪ/. Accordingly, they suggested that listeners were less sensitive to duration differences between /i/ and /ɪ/ because the spectral differences between the two vowels are sufficiently distinct. On the other hand, listeners relied on duration to greater degrees for /ɛ/-/æ/ because this vowel contrast involves a greater amount of spectral overlap. Increased overlap in the distributions of a contrast has been shown to lead to less perceptual reliance on a given dimension, presumably because more overlap makes a cue less informative (Clayards, Tanenhaus, Aslin, & Jacobs, 2008; Holt & Lotto, 2006; Nixon, 2014). Thus, increased overlap in the spectral dimensions for the /ɛ/-/æ/ contrast should lead to a weaker cue weight for that dimension for native listeners and could lead to relatively more reliance on duration. Secondly, Hillenbrand et al. (2000) also indicated that the duration differences are bigger for the /ɛ/-/æ/ contrast (i.e., approximately 84 ms) than for the /i/-/ɪ/ contrast (i.e., approximately 56 ms) on average. In fact, overlap of formant frequencies has been shown to be inversely related to durational overlap in standard British English (Bennett, 1968) as well as in North American dialects (Fridland et al., 2014). In line with the relative informativity of the two cues for the two contrasts, the native English control group in this study showed relatively more reliance on duration and less reliance on spectral cues for the /ɛ/-/æ/ contrast than for the /i/-/ɪ/ contrast. Therefore, Korean learners' relative difficulty with using spectral differences for the /ɛ/-/æ/ contrast could be attributed to the relatively greater overlap in the distribution of spectral cues for this contrast. We suggest that detailed investigation of acoustic differences and distributional information for relevant dimensions can successfully predict the relative difficulty of the acquisition of the contrasts.

4.3 The developmental stages in perceptual cue weighting

The findings of the perceptual trajectories over time from individual learners are partly in line with the hypothesized developmental stages that have been proposed for native Spanish listeners learning the English /i/-/ɪ/ contrast (Escudero, 2000). The developmental stages for acquisition of the /i/-/ɪ/ contrast in Spanish learners are recapitulated here for discussion of unexpected patterns and their relative placement within the existing stages: (0) inability to distinguish the vowel contrast, (1) use of the duration cue to distinguish between the two vowels, (2) use of both duration and spectral cues but with main reliance on duration, (3) use of spectral cues to distinguish between the two vowels.

The results of the present study suggest that L2 learners from a different language background, Korean learners of English, also broadly follow the overall sequence of hypothesized developmental stages. This finding is particularly important because this study is the first to test whether the set of developmental stages inferred from cross-sectional data based

³ We thank an anonymous reviewer for this suggestion.

on learners' varying proficiency levels corresponds to the developmental trajectories observed with longitudinal data. At the same time, however, the developmental sequences in the present study differ somewhat from those in previous studies (Escudero, 2000; Morrison, 2008). The findings of the current study indicated that learners at the initial testing time were already in different stages. This is perhaps a surprising finding for the developmental stage hypothesis (Escudero, 2000; Morrison, 2008), given that the learners had been in the country for only 2 months on average at Time 1, and had had very little or no experience with authentic spoken language input or interactions with native English speakers before arriving in Canada. It could be that some stages were skipped or that for some learners, developmental changes were too rapid to be captured in the trajectories, perhaps even happening in the first few weeks of arrival. However, it could also be that the developmental stage hypothesis does not hold up with longitudinal data, at least not for all learners. In fact, even the cross-sectional studies (Escudero, 2000; Morrison, 2008) found some individuals with very little exposure to the language who were nonetheless in advanced stages, relying on spectral cues like native speakers. In contrast to the stage hypothesis, some of our data suggest that initially, learners favor one cue or another and that they persist over time with this strategy. Those whose initial cue weighting relied on spectrum were more likely to reach native-like cue weighting by Time 4, and many of them did not pass through a duration stage at all. Our results are therefore partly consistent with previous findings, in that we see a range of strategies employed by learners. Furthermore, we demonstrated that many learners did pass through the developmental stages proposed, using duration or both duration and spectrum before establishing a pattern of spectral reliance (Escudero, 2000; Morrison, 2008). However, our results also raise the possibility that some learners begin the task of acquiring L2 vowel categories with different starting points, and these starting points may have a considerable impact on their learning trajectory.

Another key finding of this study is that the "Random" behavior with seemingly no distinction between the two vowels sometimes occurred at points along the developmental path when the learners were already showing sensitivity to a particular acoustic dimension. Thus, their behavior more likely reflects uncertainty about appropriate use of cues than insensitivity and may reflect a transition period in development. This developmental phenomenon might reflect a U-shaped learning pattern in which learners become less accurate at later stages than at earlier stages but become more accurate again at more advanced stages (Abrahamsson, 1999, 2003). Accordingly, L2 learners' apparent inability to distinguish the contrast in cross-sectional studies should be treated with caution.

It is particularly notable that Korean learners' behavior at each stage in development was mostly based on a single cue as primary. Once they noticed that the two vowels in a contrast differ, they seemed to adopt a single strategy such as relying on duration or spectral cues. Few Korean learners seemed to employ a strategy where they make use of both cues with equal reliance to categorize the contrast. Thus, learners' perceptual behavior seems to be dependent upon the most reliable cue for them at a given stage in development. This is also in line with previous research by Schertz and colleagues (Schertz, 2014; Schertz et al., 2015), which showed that cue weighting strategies used by native Korean listeners in their perception of the English stop voicing contrast were generally based on a single cue, mostly either VOT or f0 but not both. Flege et al. (1997) also observed a negative correlation between spectral and duration cue weights across learners of different L1 backgrounds and levels of experience.

4.4 Adult-child differences in L2 speech acquisition

Our findings for the adult-child differences suggest that adult learners have an initial advantage over child learners in L2 speech acquisition, at least for the /i/-/ɪ/ contrast. The adult learners made early use of spectral cues for this contrast but this was not observed among the child learners at Times 1 and 2. In terms of developmental changes over time, the group results suggest that child learners made significant improvements starting at Time 3 for both vowel contrasts, even outperforming the adults for the /ɛ/-/æ/ contrast. Although one year of immersion in an L2-speaking country is a relatively short period of time over which to observe large changes, it was nonetheless clear that the child learners significantly improved in their use

of the acoustic cues relative to the initial time point (though not all child learners showed improvement). It is especially notable that only the child learners showed improvement on the spectral dimension for the apparently more difficult /ɛ/-/æ/ contrast after one year of immersion. These findings are consistent with Aoyama et al. (2004, 2008) and Oh et al. (2011) in showing that adults have an initial advantage over children, but that over time, children show greater improvement in an L2-speaking environment. Aoyama et al. (2004, 2008) attributed children's outperformance in L2 learning to the greater amount of input they receive in an L2-speaking setting. However, in the current study, all the Korean adults attended language schools for on average 27 hours per week, providing them with L2 input comparable to that received by the children—a more controlled setting than in most previous studies. Thus, an explanation for adult-child differences based solely on the amount of L2 input is less likely than in earlier studies. It may still be the case that the adults and children under study received different kinds of input; that is, the adults might have had more non-native input than the children from their classmates in English as a second language courses (cf. Tsukada et al., 2005). However, we believe that a more plausible explanation for the adult-child differences is that the L1 and L2 phonetic systems interact more strongly in adult learners than in child learners presumably because adult learners' L1 categories are fully developed, which is likely to obstruct the formation of new categories when adult learners start to learn L2 speech sounds (Baker et al., 2008; Hazan & Barrett, 2000; Tsukada et al., 2005). Finally, it should be noted that only some children outperformed the adults even after a year. This high degree of individual variability might explain some of the inconsistent findings between children and adults in the previous literature. Furthermore, while this study was not designed to compare children of different ages, there is some evidence that the older children were more successful than the younger ones (see tables in the appendix). This fits with the results of some previous studies that similarly find that older children are more successful in L2 speech perception than younger children and adults (Shinohara & Iverson, 2013, 2015; Snow & Hoefnagel-Höhle, 1978).

4.5 Future work

This work raises several questions for future research. The results of the cue weighting patterns in the present study showed that some learners interpreted the cues in the opposite direction from native listeners (i.e., “-Spec” and “-Dur”). These patterns of cue weighting might be related to mismatches in the mapping between phonetic and lexical levels of processing. Thus, future research would benefit from examining how phonetic processing and lexical representations are related in L2 speech perception (Amengual, 2016b; Cutler, 2012; Díaz et al., 2012).

The present work also opens up possibilities for future research in adult-child differences in L2 speech perception. The findings of this study showed that some child L2 learners made greater improvement in the acquisition of L2 segments than adult L2 learners in an immersion setting where they received relatively comparable L2 input. In this situation, an explanation for adult-child differences based solely on the amount of L2 input does not seem plausible (Aoyama et al., 2004, 2008), but the differences might be the result of the differential perceptual processing of speech between adult and child learners as reflected in pre-attentive speech sound processing. Adult-child differences in L2 learning might in part be of neurobiological origin, so future studies could consider neurobiological methods, in particular, the electrophysiological method of event-related potentials, to explore neurophysiological evidence for age-related changes in pre-attentive L2 speech processing (Hisagi, Garrido-Nag, Datta, & Shafer, 2015; Ylinen, Uther, Latvala, Vepsäläinen, Iverson, Akahane-Yamada, & Näätänen, 2009). Neurobiological examination of the processing of speech sounds would add to the previous behavioral findings on L2 speech perception and also contribute to a better understanding of how the L1 and L2 interact in adult and child learners.

Perceptual weighting of speech seems to involve an interaction between linguistic aspects of speech (e.g., language-specific patterns) and non-linguistic means of processing (e.g., psychoacoustic salience). The present study has shown that most of the learners initially relied on duration cues and then a few relied on spectral cues as they gained more experience in

English. Previous work has suggested that learning duration cues might be easier than spectral cues and thus L2 learners may initially rely on duration due to its psychoacoustic salience (Bohn, 1995; Cebrian, 2006). It is interesting to note that a cue-weighting study of Dutch vowels in native Dutch children and adults (Gerrits, 2001) also found a pattern of more reliance on duration for the youngest children (4 years) shifting to an adult-like pattern for older children. This result supports the idea that duration may be privileged in the early stages of vowel contrast acquisition for both L1 and L2 learners, possibly due to its psychoacoustic salience. Future work is needed to explore the interaction of language-specific patterns of speech and psychoacoustic salience to determine their effects on cue weighting strategies (Strange & Shafer, 2008).

5.0 Conclusion

The present study has examined the developmental changes in perceptual cue weighting of English front vowel contrasts by Korean learners of English during their first year of immersion in an English-speaking environment. In doing so, this study has addressed the questions of how individuals' cue weighting strategies change over time and how cue weights are acquired across vowel contrasts which rely on the same acoustic-phonetic cues. The contribution of the present work is that the study adopted a longitudinal approach to examining developmental trajectories in the weighting of acoustic-phonetic cues by adult and child L2 learners for different vowel contrasts which, to our knowledge, had not yet been done. There were three notable findings. First, individual differences in cue weighting are not merely random variability in the learner's response patterns, but are systematically associated with the developmental trajectories of individual learners. Second, developmental trajectories vary according to vowel contrast. Finally, adult learners have an initial advantage but some child learners make greater improvements over the course of one year.

Acknowledgements

This work was supported by SSHRC grant 435-2016-0747 to Meghan Clayards and 435-2015-0490 to Heather Goad and Lydia White, and by FRQSC grant 2016-SE-188196 to Lydia White, Heather Goad and colleagues.

Appendix A

Table A1 Korean participants' demographic and language background. (LOR = length of residence in Canada at Time 1, YOE = years of English education in Korea, Hours of study = weekly hours of English study since arriving in Canada)

Group	ID	Gender	Age (years)	LOR (months)	YOE (years)	Hours of study
KA	KA01	f	39	1	6	20
	KA02	f	39	3	12	20
	KA03	f	45	4	6	25
	KA04	f	36	3	10	25
	KA05	f	40	1	10	25
	KA06	f	39	0	6	25
	KA07	f	35	0	10	15
	KA08	f	43	1	10	25
	KA09	f	35	5	10	35
	KA10	f	43	1	10	20
	KA11	f	40	1	10	15
	KA12	f	42	1	7	20
KC	KC01a	m	6	1	0.2	30
	KC01b	f	10	1	4	30
	KC02	m	11	3	4	20
	KC03	m	12	4	5	30
	KC04	m	7	3	4	30
	KC05a	f	8	1	2	40
	KC05b	m	10	1	2	40
	KC06	f	9	0	2	30
	KC07	m	8	0	1	30
	KC08	f	9	1	1	40
	KC09	f	6	5	0	25
	KC10	m	7	1	0	30
	KC11	f	11	1	5	35
	KC12	m	11	1	4	25

Appendix B

Table B1 Perceptual cue weights of individual learners based on logistic regression models for /i/-/ɪ/. Statistically significant use of acoustic cues in vowel categorization is in bold (* $p < 0.05$, ** $p < 0.01$).

ID	/i/-/ɪ/							
	Time 1		Time 2		Time 3		Time 4	
	Spectrum	Duration	Spectrum	Duration	Spectrum	Duration	Spectrum	Duration
KA01	0.93	0.00	4.74**	-2.28	-4.61**	1.46	3.10**	-1.21
KA02	0.70	0.14	1.49*	4.55**	0.71	2.44**	0.62	0.19
KA03	5.02**	0.95	3.78**	0.99*	7.09**	1.00	10.09**	0.20
KA04	3.03**	2.01**	3.03**	4.01**	6.72**	5.79**	5.92**	5.14**
KA05	-0.11	5.32**	-3.69**	-0.48	5.96**	-0.97	-4.84**	0.70
KA06	-0.15	7.66**	0.52	5.05**	0.73	10.2**	1.03	15.9**
KA07	0.00	10.03**	-0.21	3.11**	1.40	10.52**	0.24	10.84**
KA08	0.69	5.70**	0.54	1.11*	3.20**	5.14**	NA	NA
KA09	-0.18	4.17**	0.46	2.06**	-0.29	7.06**	2.28**	0.36
KA10	3.11**	4.14**	7.47**	-1.84	10.27**	-1.99	8.18**	-1.97
KA11	0.49	8.16**	0.31	8.04**	-0.40	6.49**	-0.30	7.34**
KA12	6.30**	-1.14	3.13**	1.77*	0.11	5.20**	6.46**	-0.78
KC01a	-9.94**	0.20	-5.87**	1.81	-0.47	0.61	0.43	0.15
KC01b	0.15	0.92*	0.06	-2.44**	-0.15	1.40*	0.80*	1.50*
KC02	-0.29	4.76**	1.33*	6.55**	0.57	9.00**	-1.00*	0.29
KC03	0.41	3.82**	0.19	0.65	0.48	0.95*	0.50	-0.05
KC04	-6.59**	0.27	-2.56**	1.48	1.76*	1.40*	11.53**	3.46**
KC05a	-0.43	0.84*	0.43	2.48**	0.41	5.01**	-1.14	7.17**
KC05b	0.14	3.05**	0.54	-3.57**	-0.50	1.16*	0.40	2.81**
KC06	0.29	4.66**	0.09	4.14**	-0.86	4.01**	0.71	-0.31
KC07	-0.42	-3.62**	1.60*	0.58	2.17**	0.35	0.18	4.25**
KC08	0.39	0.86*	0.19	-0.19	-0.14	-0.32	NA	NA
KC09	0.14	-0.43	-3.59**	0.49	6.98**	-1.04	4.21**	-0.40
KC10	-3.88**	-0.08	-0.61	0.36	-0.14	0.27	2.23**	-0.06
KC11	0.28	0.70	0.10	1.03*	0.61	2.77**	3.26**	4.32**
KC12	3.91**	2.23**	7.37**	1.67	6.14**	1.00	8.31**	0.85

Table B2 Perceptual cue weights of individual learners based on logistic regression models for /ε/-/æ/. Statistically significant use of acoustic cues in vowel categorization is in bold (* $p < 0.05$, ** $p < 0.01$).

ID	/ε/-/æ/							
	Time 1		Time 2		Time 3		Time 4	
	Spectrum	Duration	Spectrum	Duration	Spectrum	Duration	Spectrum	Duration
KA01	0.40	-4.83**	1.41*	3.00**	-0.16	7.78**	0.72	1.81**
KA02	0.67	-0.43	0.78	-0.05	1.45*	1.18*	0.05	-0.33
KA03	-1.10	-0.44	-0.14	-0.99*	-0.73	-1.90**	0.38	-1.83**
KA04	-0.35	-4.11**	0.29	2.05**	-0.50	4.81**	-0.06	2.27**
KA05	0.57	7.48**	-0.37	-0.60	0.91	-2.74**	0.43	-0.89*
KA06	-0.17	-8.65**	0.84	-6.95**	0.47	3.62**	0.19	-4.67**
KA07	0.17	-8.08**	-0.33	5.32**	-0.35	5.77**	0.32	12.83**
KA08	-0.25	-3.84**	0.48	5.89**	-0.1	4.66**	NA	NA
KA09	0.62	-4.14**	-0.27	6.55**	0.47	7.73**	-0.84	-1.03*
KA10	0.26	2.72**	0.00	2.96**	0.75	4.51**	2.46**	4.72**
KA11	0.18	-2.11**	-0.13	-6.31**	0.58	4.61**	-0.92	-7.61**
KA12	0.15	-3.32**	1.10	-4.87**	0.00	-11.07**	0.65	11.79**
KC01a	-0.37	-0.32	-0.37	-0.46	0.51	-0.42	-0.21	0.57
KC01b	0.00	0.15	0.15	-0.40	0.62	0.24	0.29	0.90*
KC02	0.28	0.65	-0.30	-3.40**	0.33	3.75**	-0.20	-1.35*
KC03	0.24	3.64**	0.80*	0.33	0.95*	1.09*	1.04*	0.24
KC04	-0.34	1.76*	1.58*	-0.26	5.30**	-0.44	10.89**	1.12
KC05a	-0.47	-2.91**	-0.26	-6.49**	0.14	-0.46	-0.61	7.71**
KC05b	-0.23	0.05	0.18	4.17**	-0.38	1.28*	-0.24	2.03**
KC06	-1.52	-6.28**	0.53	-8.53**	0.55	-4.30**	-0.57	-0.31
KC07	-0.54	-2.90**	-1.00	-4.73**	0.33	0.89*	-0.14	0.14
KC08	-0.19	0.29	0.42	-0.30	0.23	0.14	NA	NA
KC09	0.60	-1.12*	0.39	-0.39	-0.68	-0.30	-0.14	0.42
KC10	1.19	0.66	-3.42**	-0.99	-0.19	-0.74*	-0.99*	-0.05
KC11	0.05	-0.47	0.33	1.79*	-0.09	4.11**	3.46**	4.47**
KC12	0.05	-1.09*	1.35*	-0.15	2.31**	-0.73	13.47**	1.761

Table B3 Individual learners' developmental trajectories based on cue weighting patterns for each vowel contrast over time

ID	/i/-/ɪ/				/ɛ/-/æ/			
	Time 1	Time 2	Time 3	Time 4	Time 1	Time 2	Time 3	Time 4
KA01	Random	Spec	–Spec	Spec	–Dur	Spec+Dur	Dur	Dur
KA02	Random	Spec+Dur	Dur	Random	Random	Random	Spec+Dur	Random
KA03	Spec	Spec+Dur	Spec	Spec	Random	–Dur	–Dur	–Dur
KA04	Spec+Dur	Spec+Dur	Spec+Dur	Spec+Dur	–Dur	Dur	Dur	Dur
KA05	Dur	–Spec	Spec	–Spec	Dur	Random	–Dur	–Dur
KA06	Dur	Dur	Dur	Dur	–Dur	–Dur	Dur	–Dur
KA07	Dur	Dur	Dur	Dur	–Dur	Dur	Dur	Dur
KA08	Dur	Dur	Spec+Dur	NA	–Dur	Dur	Dur	NA
KA09	Dur	Dur	Dur	Spec	–Dur	Dur	Dur	–Dur
KA10	Spec+Dur	Spec	Spec	Spec	Dur	Dur	Dur	Spec+Dur
KA11	Dur	Dur	Dur	Dur	–Dur	–Dur	Dur	–Spec
KA12	Spec	Spec+Dur	Dur	Spec	–Dur	–Dur	–Dur	Dur
KC01a	–Spec	–Spec	Random	Random	Random	Random	Random	Random
KC01b	Dur	–Dur	Dur	Spec+Dur	Random	Random	Random	Dur
KC02	Dur	Spec+Dur	Dur	–Spec	Random	–Dur	Dur	–Dur
KC03	Dur	Random	Dur	Random	Dur	Spec	Spec+Dur	Spec
KC04	–Spec	–Spec	Spec+Dur	Spec+Dur	Dur	Spec	Spec	Spec
KC05a	Dur	Dur	Dur	Dur	–Dur	–Dur	Random	Dur
KC05b	Dur	–Dur	Dur	Dur	Random	Dur	Dur	Dur
KC06	Dur	Dur	Dur	Random	–Dur	–Dur	–Dur	Random
KC07	–Dur	Spec	Spec	Dur	–Dur	–Dur	Dur	Random
KC08	Dur	Random	Random	NA	Random	Random	Random	NA
KC09	Random	–Spec	Spec	Spec	–Dur	Random	Random	Random
KC10	–Spec	Random	Random	Spec	Random	–Spec	–Dur	–Spec
KC11	Random	Dur	Dur	Spec+Dur	Random	Dur	Dur	Spec+Dur
KC12	Spec+Dur	Spec	Spec	Spec	–Dur	Spec	Spec	Spec

References

- Abrahamsson, N. (1999). Vowel epenthesis of /sC(C)/ Onsets in Spanish/Swedish interphonology: A longitudinal case study. *Language Learning*, 49(3), 473–508.
- Abrahamsson, N. (2003). Development and recoverability of L2 codas: A longitudinal study of Chinese-Swedish interphonology. *Studies in Second Language Acquisition*, 25(3), 313–349.
- Amengual, M. (2016a). The perception and production of language-specific mid-vowel contrasts: Shifting the focus to the bilingual individual in early language input conditions. *International Journal of Bilingualism*, 20(2), 133–152.
- Amengual, M. (2016b). The perception of language-specific phonetic categories does not guarantee accurate phonological representations in the lexicon of early bilinguals. *Applied Psycholinguistics*, 37(5), 1221–1251.
- Aoyama, K., Flege, J. E., Guion-Anderson, S., Akahane-Yamada, R. R., & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics*, 32(2), 233–250.
- Aoyama, K., Guion-Anderson, S., Flege, J. E., Yamada, T., & Akahane-Yamada, R. R. (2008). The first years in an L2-speaking environment: A comparison of Japanese children and adults learning American English. *International Review of Applied Linguistics in Language Teaching*, 46(1), 61–90.
- Baker, W. W., Trofimovich, P. P., Mack, M. M., & Flege, J. E. (2002). The effect of perceived phonetic similarity on non-native sound learning by children and adults. In B. Skarabela, S. Fish, & A. H. J. Do (Eds.), *Proceedings of the 26th Boston University Conference on Language Development* (pp. 36–47). Somerville, MA: Cascadilla Press.
- Baker, W. W., Trofimovich, P. P., Flege, J. E., Mack, M. M., & Halter, R. (2008). Child-adult differences in second-language phonological learning: The role of cross-language similarity. *Language and Speech*, 51(4), 317–342.
- Baptista, B. O. (2006). Adult phonetic learning of a second language vowel system. In B. O. Baptista & M. A. Watkins (Eds.), *English with a Latin Beat: Studies in Portuguese/Spanish-English Interphonology* (pp. 19–40). Amsterdam, NLD: John Benjamins Publishing Company.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278.
- Bennett, D. C. (1968). Spectral form and duration as cues in the recognition of English and German vowels. *Language and Speech*, 11(2), 65–85.
- Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171–204). Timonium, MD: York Press.
- Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. J. Munro & O.-S. Bohn (Eds.), *Second language speech learning: The role of language experience in speech perception and production* (pp. 13–34). Amsterdam: John Benjamins.
- Bion, R. A. H., Escudero, P., Rauber, A. S., & Baptista, B. O. (2006). Category formation and the role of spectral quality in the perception and production of English front vowels. *Proceedings of the 7th Annual Conference of the International Speech Communication Association* (pp. 1363–1366). Pittsburgh, Pennsylvania.
- Boberg, C. (2010). *The English Language in Canada*. Cambridge: Cambridge University Press.
- Boersma, P., & Weenink, D. (2013). *Praat: Doing phonetics by computer* (Version 5.3.55).
- Bohn, O.-S., & Flege, J. E. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics*, 11(3), 303–328.
- Casillas, J. V. (2015). Production and perception of the /i/-/ɪ/ vowel contrast: The case of L2-dominant early learners of English. *Phonetica*, 72(2-3), 182–205.
- Cebrian, J. (2006). Experience and the use of non-native duration in L2 vowel categorization. *Journal of Phonetics*, 34(3), 372–387.

- Clayards, M., Tanenhaus, M. K., Aslin, R. N., & Jacobs, R. A. (2008). Perception of speech reflects optimal use of probabilistic speech cues. *Cognition*, 108(3), 804–809.
- Cutler, A. (2012). *Native listening: Language experience and the recognition of spoken words*. Cambridge: MIT Press.
- Darcy, I., Park, H., & Yang, C.-L. (2015). Individual differences in L2 acquisition of English phonology: The relation between cognitive abilities and phonological processing. *Learning and Individual Differences*, 40, 63–72.
- Díaz, B., Baus, C., Escera, C., Costa, A., & Sebastián-Gallés, N. (2008). Brain potentials to native phoneme discrimination reveal the origin of individual differences in learning the sounds of a second language. *Proceedings of the National Academy of Sciences*, 105(42), 16083–16088.
- Díaz, B., Mitterer, H., Broersma, M., & Sebastián-Gallés, N. (2012). Individual differences in late bilinguals' L2 phonological processes: From acoustic-phonetic analysis to lexical access. *Learning and Individual Differences*, 22(6), 680–689.
- Escudero, P. (2000). *Developmental patterns in the adult L2 acquisition of new contrasts: The acoustic cue weighting in the perception of Scottish tense/lax vowels by Spanish speakers*. M.Sc. thesis. University of Edinburgh, Edinburgh.
- Escudero, P. (2005). *Linguistic perception and second language acquisition: Explaining the attainment of optimal phonological categorization*. Ph.D. dissertation. Utrecht University, Utrecht.
- Escudero, P., Benders, T., & Lipski, S. C. (2009). Native, non-native and L2 perceptual cue weighting for Dutch vowels: The case of Dutch, German, and Spanish listeners. *Journal of Phonetics*, 37(4) 452–465.
- Escudero, P., & Vasiliev, P. (2011). Cross-language acoustic similarity predicts perceptual assimilation of Canadian English and Canadian French vowels. *Journal of the Acoustical Society of America*, 130(5), EL277–EL283.
- Francis, A. L., Kaganovich, N., & Driscoll-Huber, C. (2008). Cue-specific effects of categorization training on the relative weighting of acoustic cues to consonant voicing in English. *Journal of the Acoustical Society of America*, 124(2), 1234–1251.
- 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. E., Bohn, O.-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, 25(4) 437–470.
- Flege, J. E., & MacKay, I. R. A. (2004). Perceiving vowels in a second language. *Studies in Second Language Acquisition*, 26(1) 1–34.
- Fridland, V., Kendall, T., & Farrington, C. (2014). Durational and spectral differences in American English vowels: Dialect variation within and across regions. *Journal of the Acoustical Society of America*, 136(1), 341–349.
- Giannakopoulou A, Brown H, Clayards M, & Wonnacott E. (2017). High or low? Comparing high and low-variability phonetic training in adult and child second language learners. *PeerJ*, 5, e3209.
- Giannakopoulou, A., Uther, M., & Ylinen, S. (2013). Enhanced plasticity in spoken language acquisition for child learners: Evidence from phonetic training studies in child and adult learners of English. *Child Language Teaching and Therapy*, 29(2), 201–218.
- Gerrits, E. (2001). *The categorisation of speech sounds by adults and children*. Ph.D. dissertation, Utrecht University, Utrecht.
- Golestani, N., & Zatorre, R. J. (2009). Individual differences in the acquisition of second language phonology. *Brain and Language*, 109(2-3), 55–67.
- Hattori, K., & Iverson, P. (2009). English /r/-/l/ category assimilation by Japanese adults: Individual differences and the link to identification accuracy. *Journal of the Acoustical Society of America*, 125(1), 469–479.
- Hazan, V., & Barrett, S. (2000). The development of phonemic categorization in children aged 6–12. *Journal of Phonetics*, 28(4) 377–396.

- Heeren, W. F. L., & Schouten, M. E. H. (2010). Perceptual development of the Finnish /t-t:/ distinction in Dutch 12-year-old children: A training study. *Journal of Phonetics*, 38(4), 594–603.
- Hillenbrand, J. M., Clark, M. J., & Houde, R. A. (2000). Some effects of duration on vowel recognition. *Journal of the Acoustical Society of America*, 108(6) 3013–3022.
- Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, 97(5), 3099–3111.
- Hisagi, M. M., Garrido-Nag, K., Datta, H., & Shafer, V. L. (2015). ERP indices of vowel processing in Spanish–English bilinguals. *Bilingualism: Language and Cognition*, 18(2), 271–289.
- Holt, L. L., & Lotto, A. J. (2006). Cue weighting in auditory categorization: Implications for first and second language acquisition. *Journal of the Acoustical Society of America*, 119(5), 3059–3071.
- Ingram, J. C. L., & Park, S. G. (1997). Cross-language vowel perception and production by Japanese and Korean learners of English. *Journal of Phonetics*, 25(3) 343–370.
- Ioup, G. (2008). Exploring the role of age on the acquisition of a second language sound system. In M. Zampini & J. Hansen (Eds.), *Phonology and second language acquisition* (pp. 41–62). Amsterdam: John Benjamins.
- Iverson, P., Kuhl, P. K., Akahane-Yamada, R. R., Diesch, E., Tohkura, Y., Kettermann, A., & Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition*, 87(1), B47–57.
- Jia, G., Strange, W., Wu, Y., Collado, J., & Guan, Q. (2006). Perception and production of English vowels by Mandarin speakers: Age-related differences vary with amount of L2 exposure. *Journal of the Acoustical Society of America*, 119(2), 1118–1130.
- Kawahara, H., Takahashi, T., Morise, M., & Banno, H. (2009). Development of exploratory research tools based on TANDEM-STRAIGHT. *Proceedings of Asia-Pacific Signal and Information Processing Association, 2009 Annual Summit and Conference*. pp. 111–120.
- Kim, D., Clayards, M., & Goad, H. (2017). Individual differences in second language speech perception across tasks and contrasts: The case of English vowel contrasts by Korean learners. *Linguistics Vanguard*, 3(1).
- Kim, J.-E. (2010). Perception and production of English front vowels by Korean speakers. *Phonetics and Speech Sciences*, 2(1), 51–58.
- Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, 59(5), 1208–1221.
- Kondaurova, M. V., & Francis, A. L. (2008). The relationship between native allophonic experience with vowel duration and perception of the English tense/lax vowel contrast by Spanish and Russian listeners. *Journal of the Acoustical Society of America*, 124(6), 3959–3971.
- Kondaurova, M. V., & Francis, A. L. (2010). The role of selective attention in the acquisition of English tense and lax vowels by native Spanish listeners: Comparison of three training methods. *Journal of Phonetics*, 38(4), 569–587.
- Kong, E. J., & Edwards, J. R. (2015). Individual differences in L2 learners' perceptual cue weighting patterns. *Proceedings of the 18th International Congress of the Phonetic Sciences*, Glasgow, UK.
- Lee, J.-Y. (2009). Perception of English high vowels by Korean speakers of English. *Phonetics and Speech Sciences*, 1(4), 39–46.
- Lengeris, A. (2009). *Individual differences in second-language vowel learning*. Ph.D. dissertation. University College London, London.
- Li, D. (1998). “It’s always more difficult than you plan and imagine”: Teachers’ perceived difficulties in introducing the communicative approach in South Korea. *TESOL Quarterly*, 32(4), 677–702.

- Lipski, S. C., Escudero, P., & Benders, T. (2012). Language experience modulates weighting of acoustic cues for vowel perception: An event-related potential study. *Psychophysiology*, 49(5), 638–650.
- Mayr, R., & Escudero, P. (2010). Explaining individual variation in L2 perception: Rounded vowels in English learners of German. *Bilingualism: Language and Cognition*, 13(3), 279–297.
- Mayo, C., Scobbie, J. M., Hewlett, N., & Waters, D. (2003). The influence of phonemic awareness development on acoustic cue weighting strategies in children's speech perception. *Journal of Speech, Language, and Hearing Research*, 46(5), 1184–1196.
- Mayo, C., & Turk, A. E. (2005). The influence of spectral distinctiveness on acoustic cue weighting in children's and adults' speech perception. *Journal of the Acoustical Society of America*, 118(3), 1730–1741.
- Morrison, G. S. (2005). An appropriate metric for cue weighting in L2 speech perception: Response to Escudero and Boersma (2004). *Studies in Second Language Acquisition*, 27(4), 597–606.
- Morrison, G. S. (2007). Logistic regression modelling for first- and second- language perception data. In P. Prieto, J. Mascaró, & M.-J. Solé (Eds.), *Segmental and prosodic issues in Romance phonology* (pp. 219–236). Amsterdam: John Benjamins.
- Morrison, G. S. (2008). L1-Spanish speakers' acquisition of the English /i/-/ɪ/ contrast: Duration-based perception is not the initial developmental stage. *Language and Speech*, 51(4), 285–315.
- Morrison, G. S., & Kondaurova, M. V. (2009). Analysis of categorical response data: Use logistic regression rather than endpoint-difference scores or discriminant analysis (L). *Journal of the Acoustical Society of America*, 126(5), 2159–2162.
- Nittrouer, S., & Lowenstein, J. H. (2010). Learning to perceptually organize speech signals in native fashion. *Journal of the Acoustical Society of America*, 127(3), 1624–1635.
- Nixon, J. (2014). *Sound of mind: Electrophysiological and behavioural evidence for the role of context, variation and informativity in human speech processing*. Ph.D. dissertation. Leiden University, Leiden.
- Oh, G. E., Guion-Anderson, S., Aoyama, K., Flege, J. E., Akahane-Yamada, R., & Yamada, T. (2011). A one-year longitudinal study of English and Japanese vowel production by Japanese adults and children in an English-speaking setting. *Journal of Phonetics*, 39(2), 156–167.
- Ortega, L., & Iberri-Shea, G. (2005). Longitudinal research in second language acquisition: Recent trends and future directions. *Annual Review of Applied Linguistics*, 25, 26–45.
- Perrachione, T. K., Lee, J.-Y., Ha, L. Y. Y., Patrick C. M. Wong. (2011). Learning a novel phonological contrast depends on interactions between individual differences and training paradigm design. *Journal of the Acoustical Society of America*, 130(1), 461–472.
- R Core Team. (2008). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, URL <http://www.R-project.org>, ISBN 3-900051-07-0.
- Schertz, J. (2014). *The structure and plasticity of phonetic categories across languages and modalities*. Ph.D. dissertation. University of Arizona, Tucson.
- Schertz, J., Cho, T., Lotto, A., & Warner, N. (2015). Individual differences in phonetic cue use in production and perception of a non-native sound contrast. *Journal of Phonetics*, 52, 183–204.
- Schertz, J., Cho, T., Lotto, A., & Warner, N. (2016). Individual differences in perceptual adaptability of foreign sound categories. *Attention, Perception, & Psychophysics*, 78(1), 355–367.
- Sebastián-Gallés, N., & Baus, C. (2005). On the relationship between perception and production in L2 categories. In A. Cutler (Ed.), *Twenty-first century psycholinguistics* (pp. 279–292). Mahwah, NJ: Lawrence Erlbaum.
- Shinohara, Y., & Iverson, P. (2013). Computer-based English /r/-/l/ perceptual training for Japanese children. *Proceedings of Meetings on Acoustics*, 19(1).

- Shinohara, Y., & Iverson, P. (2015). Effects of English /r/-/l/ perceptual training on Japanese children's production. *Proceedings of the 18th International Congress of Phonetic Sciences*, Glasgow, UK.
- Snow, C. E., & Hoefnagel-Höhle, M. (1977). Age differences in the pronunciation of foreign sounds. *Language and Speech*, 20(4), 357–365.
- Snow, C. E., & Hoefnagel-Höhle, M. (1978). The critical period for language acquisition: Evidence from second language learning. *Child development*, 49(4), 1114–1128.
- Son, G. (2008). *Production and perception of English vowel categories by native Korean speakers*. M.A. thesis. Ewha Womans University, Seoul.
- Strange, W., & Shafer, V. L. (2008). Speech perception in second language learners: The re-education of selective perception. In J. G. H. Edwards & M. L. Zampini (Eds.), *Phonology and second language acquisition* (pp. 153–191). Amsterdam: John Benjamins Publishing.
- Tsukada, K., Birdsong, D., Bialystok, E., Mack, M., Sung, H., & Flege, J. E. (2005). A developmental study of English vowel production and perception by native Korean adults and children. *Journal of Phonetics*, 33(3), 263–290.
- van Leussen, J.-W., & Escudero, P. (2015). Learning to perceive and recognize a second language: The L2LP model revised. *Frontiers in Psychology*, 6.
- Wanrooij, K., Escudero, P., & Raijmakers, M. E. J. (2013). What do listeners learn from exposure to a vowel distribution? An analysis of listening strategies in distributional learning. *Journal of Phonetics*, 41(5), 307–319.
- Wong, P. C. M., & Perrachione, T. K. (2007). Learning pitch patterns in lexical identification by native English-speaking adults. *Applied Psycholinguistics*, 28(4) 565–585.
- Ylinen, S., Uther, M., Latvala, A., Vepsäläinen, S., Iverson, P., Akahane-Yamada, R., & Näätänen, R. (2009). Training the brain to weight speech cues differently: a study of Finnish second-language users of English. *Journal of Cognitive Neuroscience*, 22(6), 1319–1332.
- Yun, G. (2014). Korean listeners' perception of L2 English phoneme contrast. *Studies in Phonetics, Phonology and Morphology*, 20(2), 161–185.