Perceptual development in the categorization of pitch accent contrasts in children and adults

Kim, Jungsun1)

**ABSTRACT**

This paper examines the categorical labeling of lexical pitch accent contrasts in North Kyungsang and South Cholla Korean listeners. It focuses specifically on investigating whether the pitch accent perception of adults and children has a dialect-specific effect. To evaluate the development of perceptual identification, slopes, intercepts, and positions at categorical boundaries were computed using a logistic regression function. The results showed that differences in slopes and intercepts were significant between North Kyungsang child and adult listeners, but the same was not the case for the positions at boundaries. As far as South Cholla child and adult listeners were concerned, there was a significant difference in slopes, but not intercepts and positions at boundaries. In the present study, the comparison of intercepts and slopes at the boundaries indicated developmental differences between North Kyungsang adult and child listeners. This improvement in categorical proportion seems to be a result of developmental changes in categorical perception. For South Cholla adult and child listeners, however, perception of the non-native contrast becomes less categorical.

**Keywords:** lexical pitch accent, categorical labeling, Kyungsang, Cholla, identification task

1. Introduction

This paper investigates the properties of pitch contour in the perceptions of children and adults exposed to a Korean dialect with contrastive lexical pitch accent and one without. It especially focuses on the development of phonemic categories of pitch accent contrasts in a two-alternative forced-choice identification task. In this paper, the developmental differences with regard to lexical pitch accent between child and adult listeners will be the focus.

Regarding the development of phonemic categorization, there have been a significant amount of investigation undertaken to compare the perceptual performance of school-age children and adults (Simon & Fourcin, 1978; Elliott et al., 1982; Krause, 1982; Walley & Carrell, 1983; Gandour et al., 1986; Walley et al., 1986; Murphy et al., 1989; Kent, 1992; Ohde, Haley, & McMahon, 1996; Nittrouer & Miller, 1997; Kuhl & Meltzoff, 1997; Slawinski & Fitzgerald, 1998; Walley & Flege, 1999; Hazan & Barrett, 2000; Johnson, 2000). Perceptual identification of phonemic contrasts was carried out with children of different ages and with adults. A number of previous studies show that children and adult listeners perceive the segmental features of consonant and vowel articulation in a categorical manner. Hazan and Barrett (2000) conducted an identification task for the phonetic contrasts, /g/-/k/, /d/-/g/, /s/-/z/, and /s/-/ʃ/ for children between the ages of six and twelve years and adults and found that the children showed an increase of phonemic categorization with age that was consistent with the results obtained for the adults. In a two-alternative identification task, the identification function of the oldest group, which comprised eleven-and-a-half- to twelve-and-a-half-year-olds, was steeper than the youngest group, which comprised six- to seven-and-a-half-year-olds. Slawinski and Fitzgerald (1998) examined the identification function of a synthetic /ɹ/-/w/ by children between the ages of three and five years and adults, and found that phonemic categorization between phonemic contrasts occurred in all three groups. Here, the five-year-old children showed adult-like
perceptual performance. Burnham et al. (1991) examined the identification performance for the synthetic continuum of a native (i.e., voiced/voiceless bilabial stops) and a non-native (i.e., prevoiced/voiceless bilabial stop) sound by native English-speaking infants, two- and six-year-old children, and adults. For the identification task, the infants performed a left and right head-turn for two endpoints of a continuum. The identification of the native phonemic contrast, which involves the voiced and voiceless boundary, was categorized, again, dependent on age. With regard to the identification curves for the native phonemic contrasts, the identification performance of infants appears to be random. Children aged two years showed evidence of a sloped identification curve. Six-year-old children and adults showed obvious categorical boundaries between two native phonemic sounds. The non-native phonemic contrast involving prevoiced and voiced boundaries was poorly identified by all groups.

With regard to vowel perception, in Walley and Flege (1999), children aged between five and nine years as well as adults identified vowels in a synthetic continuum between the English vowels /ɪ/ and /iː/ in the non-word context /h_b/, and also between the English and non-English vowels /ɪ/ and /ɻ/. They showed a categorical boundary between the English /ɪ/ and /iː/ and between the English and non-English /ɪ/ and /ɻ/. The slope of the phonemic boundary for the continuum from the English /ɪ/ to the non-English /ɻ/ is less steep than that for the continuum from the English /ɪ/ to /iː/. The slopes in the identification curves by five-year-old children are less steep than those of nine-year-old children and adults. Nine-year-old children identified vowels with adult-like perceptual performance.

Generally, what these studies demonstrate, then, is that the performance of older school-age children is very similar to that of adults, and the performance of younger children can also appear to be very adult-like, though with some degradation of performance on stimuli near the categorical boundary.

Developmental differences appear in comparisons of native and other languages. The phonemic development process of the phonetic cues in a native language differs from that of other languages. Flege and Eefting (1986) examined the perception of the phonemic contrast between /θ/ and /ð/ by nine-year-old children and adults with different native language backgrounds (i.e., English and Spanish). In a categorical labeling task, English listeners showed sharper identification curves and longer voice onset time (VOT) values at categorical boundaries than Spanish listeners. The categorical boundaries of English- and Spanish-speaking children appeared at shorter VOT values than those of adult English and Spanish listeners. Regarding identification performance between different languages, Simon and Fourcin (1978) investigated the perceptual ability of English and French children aged between two and fourteen years. In the stimuli, VOT and first-formant transition were varied. As a result of the categorical boundary between the two languages, the first formant transition feature was a more important cue for English-speaking children around the age of five years than for French-speaking children. This result revealed a process involving the use of different features in identification performance between the two languages.

The research in this paper focuses on how children develop phonemic categories on lexical pitch accent (e.g., [kaci] in the North Kyungsang variety with different tonal transcriptions: HL ‘kind’, LH ‘eggplant’, HH ‘branch’) within their native dialectal background. This study employs a two-alternative forced-choice identification task to probe the relation between the phonemic categorization and dialectal variants of Korean in the perception performance of child and adult listeners.

The present study especially evaluates the identification rate of lexical pitch accent minimal pairs by non-native child and adult listeners from the South Cholla region in order to compare phonologically different prosodic patterns in dialectal variants of Korean. According to previous literature, in the South Cholla dialect, the assignment of pitch accent with high or low tone is prosodically dependent on the laryngeal property of phrasal-initial segments (Jun, 1993; Jun 1998). On the other hand, in the North Kyungsang dialect, the location of lexical pitch accent distinguishes the meaning of minimal pairs (Kim, 1976; Kim, 1988; Chung, 1991; Kim, 1997; Kenstowicz & Sohn, 1997; Jun, et al., 2006; Kim & de Jong, 2007; Kim, 2010; Kim, 2011). However, most of the previous studies have not investigated the developmental patterns of lexical pitch accent from children to adults (Kim, 1976; Kim, 1988; Chung, 1991; Kim, 1997; Kenstowicz & Sohn, 1997; Jun, et al., 2006) in the field of phonetics and phonology. Recently, Kim & de Jong (2007) investigated the mimicry responses in the light of the relation between perceptual identification and production. This research employed the mimicry paradigm and was conducted with adult participants from two dialectal regions. The result implied that the acquisition of lexical pitch accent occurs at the early stages of speech perception and production. Kim (2010) reported the acoustic characteristics of lexical pitch accent produced by child and adult speakers in a Korean cross-dialect, suggesting the occurrence of developmental changes in the process of
phonological acquisition, such as the difference of f0 (i.e., fundamental frequency) values on pitch contours of pitch accent minimal pairs. In this paper, to find out whether lexical pitch accent is perceived in the process of phonological acquisition, a two-alternative forced-choice identification task is conducted for child and adult listeners from two dialectal regions.

Previous studies on the pitch accent system of Korean did not investigate the development of a prosodic category by children. Research on the prosodic category in the process of phonological acquisition is necessary for explaining how a lexical item as a prosodic unit is encoded in the developmental patterns between adult and child participants. From recent studies on lexical pitch accent (Kim & de Jong, 2007; Kim, 2010; Kim, 2011), the prosodic categories of lexical pitch accent were found to be distinguished by native dialectal speakers (e.g., North Kyungsang) but not by non-native dialectal speakers (e.g., South Cholla). On the basis of a study on developmental differences, the prosodic categories that are perceived by children will shed some light on the prosodic development of children. Based on phonological development in cross-linguistic studies, child participants are expected to show the development of phonemic categorization through the acquisition of tacit knowledge of a local phonology. In phonological acquisition, previous studies on cross-language perception have reported that children from the ages of six or seven years show indications of adult-like perceptual performance (Krause, 1982; Gandour et al., 1986; Burnham et al., 1991; Ohde, Haley, & McMahon, 1998; Walley & Flege, 1999). Studies on lexical pitch accent in Korean have not examined the perceptual abilities acquired by child listeners from their local dialects. The research on lexical pitch accent of child listeners will be helpful for understanding the prosodic system of Korean.

2. Experimental Method

2.1 Participants

Two groups of participants were recruited, one from the central region (i.e., Daegu) speaking North Kyungsang Korean with a pitch accent distinction, and the other from the central region (i.e., Kwangju) speaking South Cholla Korean without a pitch accent distinction. The current study examines whether six- and seven-year-old children’s perceptual capabilities are representative of speakers of their local dialect. Five child participants aged between six and seven years were recruited from the two above-mentioned regions, that is, five from each region. To compare the perceptual performance between child and adult participants, twelve adult participants from the North Kyungsang region and ten from the South Cholla region were recruited. All children and adults were natives, and all of them were compensated financially. None of them reported any kind of hearing impairment.

2.2 Stimuli

Stimuli were created by a pitch-synchronous overlap and add (PSOLA) algorithm using a Praat 4.4 (Boersma & Weenink, 1992-2009). A nine-step pitch accent continuum was generated from natural tokens (e.g., [moi]: HL vs. LH ‘feed’, ‘conspiracy’, [moɾɛ]: HL vs. HH ‘sand’, ‘the day after tomorrow’, [yanŋ]: LH vs. HH ‘wool’, ‘adoptive mother’) produced by a North Kyungsang speaker. Three pitch accent continua corresponding to HL-LH, HL-HH, and HH-LH were created. In a two-alternative forced-choice task, three nine-step continua were employed for the adult and child listeners. To make the protocol appropriate to the children’s abilities, stimuli were generated for only three continua (e.g., HL→LH, HL→HH, HH→LH).

2.3 Procedure

The experiment included a two-alternative forced-choice identification task. Before carrying out the task, the participants were shown illustrations of the target words to help them understand the words to be used in the experiment. To check the familiarity of the target words, the child participants were asked what the pictures are describing. The child participants consistently provided answers for the pictures that corresponded to the target words.

In the identification task, each stimulus for each continuum was presented over high-quality headphones in a quiet room. This experiment included 216 trials containing 3 blocks of 72 trials, and the trials within each block were presented in random order, using perception software ALVIN (Hillenbrand & Gayvert, 2005). Each stimulus block was preceded by 32 practice trials. The participants were asked to press one of the buttons indicating the target words. For the comparison with the adult speakers, the identification scores obtained on the 216 trials were used.

2.4 Data analysis

For the analysis of identification functions by adult and child speakers from the North Kyungsang and South Cholla regions, a logistic regression function was computed with the formula, \[ Y = \frac{1}{1+\exp(-(a+bx))} \]. The categorical boundary was derived with the
formula, $x = -a/b$. The regression coefficients (a, b) corresponding to the intercept and slope were estimated using binary logistic regression. Two-alternative forced-choice labeling responses were evaluated using the regression coefficients (i.e., the intercept and slope of the identification response curve) and positions of the categorical boundary (i.e., the 50% cross-over midpoint of the identification response curve). The logistic regression function predicts the probability of occurrence of an event (i.e., the identification rate), on the basis of observations of whether or not the event did occur. The function that fits the data in logistic regression can be conceptualized as a line that can be described mathematically with the help of the formula, $Y = 1/(1+\exp(-a+bx))$. The intercept of the line represents its position in geometric space, and the slope represents the shape of the line. In other words, the value of intercept is the y-intercept at which the line crosses the y-axis when x is equal to zero, and the value of slope indicates the extent of increase in the line when x increases by one. Considering the logistic regression function on the categorization of pitch accent minimal pairs, it is predicted that the higher the values of intercept, the lower the values of slope, thus increasing the steepness in the probability curves. To describe the difference and similarity in the logistic regression function for each group and each pitch accent pattern, the present study conducted an independent $t$-test and a one-way analysis of variance (ANOVA) for the values of the regression coefficients and positions of the categorical boundary.

3. Results

Figure 1 summarizes the values of slope and intercept and the position of the categorical boundary for logistic fits to the identification functions for the Kyungsang and Cholla dialect groups for the three lexical pitch accent minimal pairs.

In Figure 1, on comparing the two dialect groups, North Kyungsang and South Cholla child listeners, the intercept was found to be significantly different for all three pairs: HL-LH ($t (4) = 3.891, p = .017$), HL-HH ($t (5) = 4.959, p = .004$), and HH-LH ($t (8) = 2.808, p = .023$). Identification slopes were also significantly different between the two groups for all three: HL-LH ($t (5) = -3.271, p = .021$), HL-HH ($t (4) = -5.255, p = .005$), and HH-LH ($t (8) = -2.856, p = .021$). The position of the categorical boundary was significantly different for HL-LH ($t (7) = 2.479, p = .042$), but not for HL-HH ($p = .314$) and HH-LH ($p = .454$). The categorical boundaries for North Kyungsang listeners are centered around stimulus number 5, but for South Cholla listeners, the 50% identification cross-over estimate appears around stimuli numbers 1, 2, and 3, though this cross-over point is essentially meaningless owing to the very flat categorization functions in the South Cholla listeners. Thus, the categorical boundary for the North Kyungsang child listeners is much sharper than that for the South Cholla listeners. Categorization in perceptual identification is exhibited by the North Kyungsang child listeners.

To distinguish the three pitch accent minimal pairs, HL-LH, HL-HH, and HH-LH, a one-way ANOVA was conducted for North Kyungsang and South Cholla child listeners. For North Kyungsang listeners, the ANOVA showed that the three pitch accent minimal pairs were not significantly different for the intercept ($p = .622$), slope ($p = .563$), or position of the categorical boundary ($p = .902$). For the South Cholla listeners, the ANOVA also showed that the three pitch accent patterns were not significantly different for the intercept ($p = .345$), slope ($p = .168$), or position of the categorical boundary ($p = .752$). There was no difference in the three pitch accent pairs from each dialectal group. This shows that, regardless of which accent pattern contrast is being examined, having lexical contrast is a critical factor for perceptual identifications.

Figures 2, 3, and 4 exhibit box plots of intercepts, slopes, and positions of the categorical boundary, respectively, for all listeners by age and dialect group. For the North Kyungsang child and adult listeners, the intercept ($t (83) = -4.403, p < .001$) and slope
(t (83) = 4.369, p < .001) were significantly different. However, this was not the case for the position of the categorical boundary (p = .208). The identification function was steeper for adults than for children. The intercept and slope indicate a developmental difference, but the position of the categorical boundary shows similarity between child and adult listeners. On the other hand, for the South Cholla child and adult listeners, the intercept (p = .613) and the position of the categorical boundary (p = .084) were not different, but the slope was significantly different (t (73) = 5.072, p < .001).

4. Discussion

This paper focuses on the perception of pitch accent contrast by listeners of different dialects. The results indicated that the intercept, slope, and position of the categorical boundary of the identification functions were significantly different between North Kyungsang and South Cholla listeners. The identification curve of the North Kyungsang child listeners with a lexical pitch accent dialect was significantly steep, but the curve for the South Cholla child listeners was rather flat. Overall, the identification functions of the South Cholla child listeners were random and showed much variability among individuals.

4.1 Mean identification responses

Figure 1 shows that the identification curves for the HL-LH, HL-HH, and HH-LH patterns are much steeper for the North Kyungsang listeners than for the South Cholla listeners. The responses of the North Kyungsang listeners are almost identical across all pairs. For the South Cholla listeners, the identification curves between each pitch accent minimal pair are not very steep (nearly flat in two of the continua). In addition, the identification curves between the pitch accent patterns of the South Cholla listeners appear as flat lines, as compared to the responses of the North Kyungsang listeners. The HH-LH pattern of the South Cholla child listeners may show an inclination toward forming an identification curve, which is the reverse of what happens with the HL-LH and HL-HH pitch accent patterns.
4.2 Identification responses of child and adult listeners

Figures 1, 2, 3, and 4 show developmental differences in lexical pitch accent between child and adult listeners from the North Kyungsang and South Cholla regions. The identification functions point out differences in the lexical status of native and non-native speakers as well as developmental differences. There is a difference in the intercept and slope between the North Kyungsang child and adult listeners, while the opposite is true for the position of the categorical boundary. The intercept is higher and the slope is steeper for the North Kyungsang adult listeners than for the North Kyungsang child listeners. The intercept and slope play an important role in distinguishing the perceptual identifications shown by the child and adult listeners, but the position of the categorical boundary does not reveal any difference between the two groups. The difference in the intercept and slope between the North Kyungsang child and adult listeners seems to indicate the developmental effect of perception on lexical pitch accent. The language patterns of the children are developing in the same manner, but they are less proficient, thus yielding a flatter identification function.

Specifically, considering Figures 1 and 3, the slopes of the identification functions of the North Kyungsang listeners become progressively steeper for the continuum of native dialect, and the responses of the child listeners in the identification curves are not as obvious as those of the adult listeners. In the continuum of the HL-LH pattern, progressing from a falling to a rising contour, the slope of the identification function for the child and adult listeners is the steepest. For the continua progressing from a falling or rising contour to a high tone (i.e., HH), the manner of identification of the child listeners is not the same as that of the adult listeners.

Further, the identification curves of the South Cholla listeners are shallower than those of the North Kyungsang listeners. The South Cholla child and adult listeners show random patterns for the intercept and position of the categorical boundary, though the slope of identification functions of the South Cholla adult listeners is steeper than that of the child listeners. Developmental differences are dependent on how sharply phonetic categories are aligned on variations in lexical knowledge of native sounds.

Accordingly, the perceptual identification patterns of the North Kyungsang child listeners are different compared with the South Cholla child listeners. In other words, the North Kyungsang child listeners show a tendency toward perceptual categorization for each pitch accent pattern, but the identification curves of the South Cholla child listeners are very flat. By comparing the South Cholla child and adult listeners, the slope of the identification curves is found to be different, but the intercept is not.

4.3 Individual identification responses of child listeners

The individual identification functions of the child listeners in Figure 5 show that their perception is similar to that of adult listeners within one dialect, but is different across dialects. The perceptual identifications by the child listeners exhibit different phonological patterns from both dialectal regions.

Across most of the North Kyungsang child listeners, the intercept and slope do not show great variation, and the identification curves are generally sharply categorical, though for the HL-HH and HH-LH patterns, the response for one of the participants is less categorical. The intercept is relatively different. For the HL-LH pitch accent pattern, the identification functions are categorical because the pitch fall is significantly different for a falling contour for the HL pattern and a rising contour for the LH pattern. Every North Kyungsang child listener shows a lesser tendency toward categorizing the HL-HH pair.

The identification curves of the South Cholla child listeners appear randomly. The intercept and slope are greatly varied depending on the individuals. The categorical boundaries across the responses of all the listeners are inconsistent. The identification curves are reversed at the endpoint of each minimal pair for one listener each in the HL-LH and HL-HH patterns and four listeners in the HH-LH pattern. In most of the cases, the identification curves are not steep, and appear as relatively flat lines.

![Identification curves of lexical pitch accent by child listeners from North Kyungsang and South Cholla regions](image-url)
The individual identification functions are quite consistent across the North Kyungsang child listeners, suggesting that the categorical boundaries for most of the individual listeners are sharp. Among the South Cholla child listeners, the identification functions show substantial variability across individual listeners. In addition, the individual identification curves appear as flat lines. Thus, the identification functions of North Kyungsang and South Cholla child listeners show different patterns. The categorical boundaries in perceptual identification are sharper for the North Kyungsang child listeners.

5. Conclusion

The developmental process of lexical pitch accent in North Kyungsang child listeners was examined by comparing them with South Cholla child listeners. South Cholla adult listeners do not have the lexical constraints of lexical pitch accent as do North Kyungsang listeners. The properties of the pitch accent perceptions of North Kyungsang listeners are affected by the phonology of the local dialect, and the lexical patterns of the North Kyungsang child listeners are tracked by the North Kyungsang adult listeners. The results of this research imply that adult-like phonemic categorization is achieved from the early linguistic experience of a local dialect and may be largely in place by the age of six years. The phonological property is lexicalized through a continuous learning process within the environment of the local community, and there is some evidence that this learning process is still evident in perceptual behavior through early school years. However, in order to be able to generalize the results of this study in a more comprehensible manner, it is necessary to perform a longitudinal study using a larger sample consisting of children and to employ other methods of analyzing perception.

References


Kim, Jungsun
Department of English Language and Literature
Yeungnam University
214-1 Dae-Dong, Gyungsan, Gyungsangbuk-Do, Korea 712-749
Email: jngsnkim@gmail.com