CUE WEIGHTING OF STATIC AND DYNAMIC VOWEL PROPERTIES BY ADULTS AND CHILDREN WITH NORMAL LANGUAGE AND SPECIFIC LANGUAGE IMPAIRMENT

Ralph Ohde & Stephen Camarata

Dept. of Hearing and Speech Sciences, Vanderbilt University,
Nashville, TN, USA

ralph.n.ohde@vanderbilt.edu

ABSTRACT

The purpose of this study was to determine whether children with and without language impairment give more perceptual weight than do adults to dynamic spectral cues versus static cues, when identifying vowel sounds. Three experimental stimulus sets were presented, each with 30 ms stimuli. The first consisted of unchanging formant onset frequencies ranging in value from frequencies for [i] to those for [α], corresponding to a bilabial stop consonant. The second two consisted of either an [i] or [α] onset frequency with a 25 ms portion of a formant transition whose trajectory was toward one of a series of target frequencies ranging from those for [i] to those for [α]. Ten typically developing children between the ages of 3;8 and 4;1, seven children with specific language impairment (SLI) between the ages of 5;1 and 6;11, and a control group of 10 adults identified each stimulus as [bi] or [bα]. The results showed developmental effects: the typically developing children relied more heavily than the adults did on the static formant onset frequency cue to identify the vowels. The SLI children followed one of two perceptual profile patterns: 1) nonrandom perception of [α] onset stimuli, but random perception of [i] onset stimuli; and 2) nonrandom perception of both [i] onset and [α] onset stimuli, but phonetic boundaries different from typically developing children and adults. The results will be discussed in regard to the Developmental Perceptual Weighting Shift theory, and perceptual profile differences in SLI children.

INTRODUCTION

Vowel perception, the subjective identification of a spoken vowel sound by a listener, incorporates properties of not only the syllable nucleus, or "quasi-steady state" portion, of the vowel, but also dynamic properties such as intrinsic vowel durational differences and the formant transitions into and out of the vowel (Jenkins, Strange, & Edman, 1983; Nearey, 1989; Strange, Jenkins, & Johnson, 1983; Strange, 1989a). Typically developing children's perception of speech appears to be different from that of adults, probably a result of both developmental immaturity in auditory perception and differences in which part of the signal they attend to; the various influence of these two factors is debated (Parnell & Amerman, 1978). While adults tend to give perceptual weight to both static and dynamic acoustic properties of speech (Strange, 1989b), some researchers have speculated that children appear to give more weight than adults do to the dynamic spectrally changing properties of the signal (Nittrouer, 1992). Although there currently exists only a handful of studies looking specifically at vowel perception, various studies have demonstrated differences in perception of consonants and vowels between children and adults (e.g., Parnell & Amerman, 1978; Murphy, Shea, & Aslin, 1989; Nittrouer, 1996; Ohde & Haley, 1997; Ohde *et al.*, 1996), concluding that young children appear to use dynamic cues

such as formant transitions relatively more than do adults in phoneme perception. Nitttrouer (1996) terms this perceptual process the Developmental Weighting Shift. One recent study by Sussman (2001) did appear to contradict Nitttrouer's (1996) developmental weighting shift theory, finding that children actually gave more perceptual weight to the static than dynamic vowel properties. However, she acknowledges that this may have been due to children's immature perceptual systems; they may have attended to the longer steady-state information than formant transitions because of its saliency only rather than because of a specific weighting strategy.

In the study by Sussman (2001), vowel perception in children with SLI was also examined. The examination of speech perception strategies in SLI children is important, since previous research has indicated that these children may display deficits in the ability to process rapidly changing acoustic signals such as formant transitions (Leonard, 1998; Tallal, Stark, & Mellits, 1985a,b). Sussman found that SLI children were significantly poorer than adults and typically developing children in the identification of the [i] vowel from either formant transitions or vowel steady states. In reviewing the literature, it becomes clear that very little research has investigated vowel perception in children. Developmental research is contradictory regarding the salience of static versus dynamic cues in vowel perception (Ohde *et al.*, 1996; Ohde & Haley, 1997; Sussman, 2001). Moreover, little research on vowel perception exists in children with SLI compared to typically developing children. Thus, further investigation is necessary to examine the role of these cues in the development of vowel perception in children with and without SLI, and to determine why it might occur in children as compared to adults.

METHOD

Test Stimuli

The stimuli were generated using the Klatt cascade/parallel formant synthesizer (Klatt & Klatt, 1990) at a sampling rate of 10 kHz. The output was low-pass filtered with a cutoff frequency of 4 kHz. Two vowels, [i] and [ɑ], were used as the bases for the construction of stimuli. These two vowels were chosen to be very different in terms of their first (F1) and second (F2) formant frequencies (onset and target frequencies). The frequencies were based on syllables with a bilabial stop-consonant onset. The stimuli were synthesized with source and resonance parameters appropriate for an adult male vocal tract. Five different pairs of control stimuli and three experimental continua were generated. The first pair of control stimuli was naturally produced 300 ms syllables [bi] and [bɑ]. The second pair was synthesized 300 ms full-length [bi] and [bɑ] syllables. The last three pairs were "practice stimuli" of 30 ms duration, and represented the endpoint stimuli from the experimental stimulus continua. The three conditions of 30 ms experimental stimuli each contained the following F2 properties: (1) F2 static-onset only; (2) F2 [i]-onset + transition; and (3) F2 [ɑ]-onset + transition. The first contained static information only, and the second and third contained dynamic and static information. These three spectral conditions will be referred to as "static-onset", "[i]-onset", and "[ɑ]-onset". In the static-onset condition, F2 values varied in 10 equal steps of 100 Hz from endpoint [i] to endpoint [ɑ]. The two dynamic (F2 onset + transition) conditions contained an initial F2 onset frequency (5 ms) appropriate for a labial stop in the context of either an [i] or an [ɑ] vowel and a F2 transition toward a target frequency which was not reached (only 25 ms of the typical 45 ms

transition duration was presented). The F2 onset was held constant at either 1800 Hz for the F2 "[i]-onset" condition or 900 Hz for the F2 "[α]-onset" condition, and the F2 transition varied in 11 steps that approached F2 target frequencies ranging from canonical [i] (2200 Hz) to canonical [α] (1200 Hz).

Participants

Ten typically developing children ages 3;8 to 4;1 (mean 3;11), seven SLI children ages 5;1 to 6;11 (mean 6;3), and ten adults ages 23;10 to 32;0 (mean 25;11) participated in this study. The typically developing children exhibited age-appropriate speech and language scores on the *Arizona Articulation Proficiency Scale, Third Revision* (Fudala, 2000) and the *Test of Early Language Development, 2nd ed.* (Hresko et al., 1991). The SLI children scored at least one standard deviation below the mean for their age on at least one of the standardized language tests [TACL (Carrow-Woolfolk, 1985) or CELF-P (Wiig et al., 1992): linguistic concepts or recalling sentences subtests]. The adults used a standard American dialect and exhibited no evidence of speech or language impairment based on informal assessment. All subjects were native speakers of English and passed a pure-tone hearing screening at 20 dB for octave frequencies from 500-4000 Hz prior to each testing session.

Procedure

The three experimental stimulus conditions were presented separately to each participant in a 2-alternative forced-choice task. Participants were asked to listen to each stimulus and decide whether the syllable they heard was [bi] or [bα]. Participants indicated their choice by pointing to one of two puppets for children, or by pressing buttons on a response box marked "[i]" and "[α]" for adults. The adults and children received practice in making these choices, using the control stimuli. Each stimulus set contained four presentations of each stimulus as well as four presentations of each full-length syllable, all in randomized order.

Statistical Analyses

A repeated measures analyses of variance on category boundaries and slopes was performed on the data for the [α]-onset only (the slope results are not reported in this paper). This analysis included the between factor of group (adult, three-year-old, and SLI). The phonetic boundaries (category boundaries), which mark the 50% point between sound categories, were determined by assigning to each stimulus a number ranging from either 1 to 11 and performing least-mean-squares analysis of the Z transformations of the percent identification data (Ferguson, 1976). Significant results at the $p < 0.05$ or $p < 0.01$ levels are reported. Statistical details are reported only for significant effects. In addition, ANOVAs were used to analyze the stimulus at each category end (e.g. stimuli 1 = [i]; stimuli 11 = [α]) for the static-onset, [i]-onset, and [α]-onset conditions. The ANOVAs were performed on the arcsine transformed percents for the between factor of group (adult, 3-year-old, and SLI) and the within factor of stimulus.

RESULTS AND DISCUSSION

In perceiving speech, several processes occur including phonetic boundary categorization and phonetic identification. Thus, as listeners we must be capable of identifying phonemes, and we must also be able to categorize sounds ranging from one phoneme to another phoneme. Across the three stimulus conditions, phonetic boundaries for the SLI group were obtainable for

only the [α]-onset stimuli, whereas typically developing 3-year-olds and adults categorized all three stimulus conditions. Thus, SLI children appear to have a basic deficit in phonetic boundary categorization. As illustrated in Figure 1, a repeated measures analysis of variance of the phonetic boundaries for the [α]-onset stimuli revealed a significant group effect [$F(2,22)=8.17$; $p<0.01$]. Follow-up comparisons showed that the three groups were significantly ($p<0.05$) different from each other. These results reflect two distinct processes. The difference between typically developing 3-year-old children and adults reveal a developmental effect, and the difference between the 3-year-old children and SLI children indicate an effect of language disorder. The mean phonetic boundaries were 6.726, 5.747, and 4.832 for the adults, 3-year-old children, and SLI children, respectively. These boundaries show that, as you go from adult to 3-year-old to SLI groups, fewer of the vowels in the [α]-onset condition were perceived as [i].

Figure 1. Mean percent [i] responses of typically developing and SLI children, and adults for the [α]-onset condition.

One prediction based on the above phonetic boundary results would be that SLI children will have a greater difficulty in the phonetic identification of [i] than [α]. As a test of this prediction, ANOVAs were used to analyze the stimulus at each category end (e.g. stimulus 1 = [i]; stimulus 11 = [α]) for the static-onset, [i]-onset, and [α]-onset conditions. The ANOVAs were performed on the arcsine transformed percents for the between factor of group (adult, 3-year-old, and SLI) and the within factor of stimulus (e.g. stimulus 1 and stimulus 11). The results revealed strong

support for this prediction. For each condition, there was a significant group X stimulus interaction (static onset: $F(2,24) = 7.14$; $p < 0.01$]; [i]-onset: $F(2,24) = 29.56$; $p < 0.01$]; [α]-onset: $F(2,24) = 9.42$; $p < 0.01$]. Follow-up comparisons across all three conditions showed that adult and 3-year-old child groups were significantly more accurate in the identification of [i] than the SLI group. On the other hand, with the exception of the [i]-onset condition, all groups identified the most [α]-like stimulus in the experimental continuum the same. In the [i]-onset condition, follow-up comparisons revealed that adult and 3-year-old child groups were significantly more accurate in the identification [α] than the SLI group.

CONCLUSIONS

1. SLI children have perceptual deficits in both the placement of phonetic boundaries (the 50% point separating two vowel sound categories) and in phonetic identification.
2. For SLI children, the deficits in perception appear to be related to acoustic correlates of vowels in that perception of the diffuse vowel [i] was much poorer than the grave vowel [α].
3. The poorer identification of [i] by SLI children may relate to a deficit in auditory processing of high frequency energy characteristic of diffuse vowels.
4. Contrary to the developmental weighting shift hypothesis, the typically developing and SLI child groups appeared to rely more heavily than the adult group on the static-onset cue information, while adults appeared to give more equal weight to static and dynamic information.

ACKNOWLEDGEMENTS

This research was supported in part by NIH Grants DC00523-08, P30 HD15052, and P50 DC03282. The authors thank Mary Camarata, Megan Duncan, Courtney Jacks, and Susan Usery for assistance on this project. We extend our sincere thanks to the adults and the children and their parents who participated in this research, without whose cooperation, help, and patience such study could not have been completed.

REFERENCES

- Carrow-Woolfolk, E. (1985). *Test for Auditory Comprehension of Language*, Allen, TX: DLM Teaching Resources.
- Ferguson, G.A. (1976). *Statistical Analyses in Psychology and Education*, New York: McGraw-Hill.
- Fudala, J.B. (2000). *Arizona Articulation Proficiency Scale, Third Revision*, Los Angeles, CA: Western Psychological Services.
- Hresko, W.P., Reid, D.K. & Hammill, D.D. (1991). *Test of Early Language Development-2nd ed*, Austin, TX: Pro-Ed.
- Jenkins, J.J., Strange, W. & Edman, T.R. (1983). Identification of vowels in 'vowelless' syllables, *Perception and Psychophysics*. 34, 441-450.

- Klatt, D. H., & Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers, *Journal of the Acoustical Society of America*, 87, 820-857.
- Leonard, L. (1998). *Children with Specific Language Impairment*, Cambridge, MA: MIT Press.
- Murphy, W.D., Shea, S.L. & Aslin, R.N. (1989). Identification of vowels in 'vowelless' syllables by 3-year-olds, *Perception and Psychophysics*, 46, 375-383.
- Nearey, T.M. (1989). Static, dynamic, and relational properties in vowel perception, *Journal of the Acoustical Society of America*, 85, 2088-2113.
- Nittrouer, S. (1992). Age-related differences in perceptual effects of formant transitions within syllables and across syllable boundaries, *Journal of Phonetics*, 20, 351-382.
- Nittrouer, S. (1996). Discriminability and perceptual weighting of some acoustic cues to speech perception by 3-year-olds, *Journal of Speech and Hearing Research*, 39, 278-297.
- Ohde, R.N. & Haley, K.L. (1997). Stop consonant and vowel perception in 3- and 4-year-old children, *Journal of the Acoustical Society of America*, 102, 3711-3722.
- Ohde, R.N., Haley, K.L. & McMahon, C.W. (1996). A developmental study of vowel perception from brief synthetic consonant-vowel syllables, *Journal of the Acoustical Society of America*, 100, 3813-3824.
- Parnell, M.M. & Amerman, J. (1978). Maturational influences on perception of coarticulatory effects, *Journal of Speech and Hearing Research*, 21, 682-701.
- Strange, W. (1989a). Dynamic specification of coarticulated vowels spoken in sentence context, *Journal of the Acoustical Society of America*, 85, 2135-2153.
- Strange, W. (1989b). Evolving theories of vowel perception, *Journal of the Acoustical Society of America*, 85, 2081-2087.
- Strange, W., Jenkins, J.J. & Johnson, T.L. (1983). Dynamic specification of coarticulated vowels, *Journal of the Acoustical Society of America*, 68, 1622-1625.
- Sussman, J.E. (2001). Vowel perception by adults and children with normal language and specific language impairment: Based on steady states or transitions? *Journal of the Acoustical Society of America*, 109, 1173-1180.
- Tallal, P., Stark, R., & Mellits, D. (1985a). The relationship between auditory temporal analysis and receptive language development: Evidence from studies of developmental language disorder, *Neuropsychologia*, 23, 527-534.
- Tallal, P., Stark, R. & Mellits, D. (1985b). Identification of language-impaired children on the basis of rapid perception and production skills, *Brain and Language*, 25, 314-322.
- Wiig, E.H., Secord, W. & Semel, E. (1992). *Clinical Evaluation of Language Fundamentals-Preschool*, San Antonio, TX: The Psychological Corporation, Harcourt Brace Jovanovich.