CHILDREN'S CLEAR SPEECH SUGGESTS WORD-LEVEL TARGETS: PRELIMINARY EVIDENCE

Melissa A. Redford and Christina E. Gildersleeve-Neumann

The University of Oregon and Portland State University
redford@uoregon.edu and cegn@pdx.edu

ABSTRACT

The development of clear speech was examined in a cross-sectional study of preschool children aged 3, 4, and 5 years old. Thirty children produced target monosyllabic words with monophthongal vowels in clear and casual speech conditions. Vowel acoustics were measured and adults were asked to provide clear speech ratings on either the vowel or the whole word. The results provided little evidence that young children hyperarticulate vowels in clear speech. Rather, the results suggest that children aim for more adult-like word targets in clear compared to casual speech.

Keywords: Speech development, speech styles, speech targets, vowel acoustics.

1. INTRODUCTION

Studies on the acoustics of clear speech in adults indicate that speakers achieve clearer, more intelligible speech when they produce more extreme articulations of segmental targets. For example, Picheny *et al.* [9], Moon & Lindblom [8], and Bradlow *et al.* [2] have all found larger vowel spaces in clear speech compared to citation or lower intelligibility speech. Johnson *et al.* [3] labeled the perceptual equivalent of this result the Hyperspace Effect, and suggested that it emerges from the representation of hyperarticulated phonetic targets; specifically, targets that highlight phonemic distinctiveness. Lindblom [5] has also argued that clear speech is aimed at paradigmatic distinctiveness, presumably to facilitate speech decoding.

The hypothesis of paradigmatic distinctiveness assumes a speaker whose actions are guided by an abstract phonology. If this phonology develops slowly from lexical representations, as Lindblom [6] and others have suggested (e.g., [1], [7]), then the distinctiveness hypothesis predicts that young children will not control a clear speech style until they have represented the paradigmatic structure of their language. An alternative prediction, consistent with the view that phonology emerges from the lexicon, is that young children may style shift with lexical rather than featural targets in mind. The current study examined the development of clear speech strategies with these different ideas in mind.

2. METHODS

2.1. Speech Sample

Three groups of 10 American-English speaking preschool children participated in the production study (N = 30). Children ranged in age from 3;4 to 3;7, from 4;4 to 4;7, and from 5;3 to 5;8.

2.1.1. Stimuli

Stimuli were pictures of familiar objects (e.g., tree, pig, bed, cat, bus, dog, drum, juice) obtained from Boardmaker (Johnson-Mayer, Inc.), ensuring that all pictures had been previously tested for ease of recognition. Although sixty-three pictures were selected to broadly sample children's speech sound repertoires, the current study focused on the subset of pictured objects with monosyllabic names and had monophthongal syllabic nuclei (N = 25).

2.1.2. Recordings

The preschool children participated in two picture naming tasks designed to elicit clear and casual speech, respectively. In the clear speech condition, the experimenter asked the child whether or not she could name each of the objects shown on 5 × 7 inch laminated cards. Children were recorded as they named the pictures, which were presented one at a time in a randomized order. To obtain maximally clear speech, children were frequently reminded to use their "big girl" or "big boy" voice when naming pictures.

In the casual speech condition, the experimenter left the room so that the child and parent could play together. Before she left, the experimenter provided the dyad with a set of wooden blocks. Each block had a stimulus object pictured on two sides. Parents were instructed to play a game that would encourage children to spontaneously picture name. The dyad was recorded while playing.

The subset of pictured words produced in the casual speech condition were matched to their clear speech counterparts. That is, speech style was a within-subjects factor. Since children varied in their volubility and named different objects in the casual speech condition, word type was treated as a between-subjects factor, characterized by vowel height, front-backness, or syllable structure in the analyses.

2.1.3. Acoustic measures

The 30 preschool children produced a total of 544 clear and casual monosyllabic words with monophthongal vowel nuclei (i.e., 272 matched pairs). Acoustic measures focused on the vowel. The recorded words were displayed as oscillograms and spectrograms simultaneously. Vowel onsets and offsets were identified from amplitude and frequency changes in the periodic waveform. Auditory judgments confirmed visual segmentations. Acoustic duration was recorded, and three formant measures were taken: one at F1 midpoint, one at F2 onset, and one at F2 midpoint. The formant measures were made from visual inspection of the spectrogram, and were checked against spectral slices of the waveform. Obtained values were consistent with previously published data on American-English speaking children's vowels (e.g., [4]).

2.2. Ratings

Two groups of 10 university undergraduates (N = 20) provided clear speech ratings on the 544 words. The rating task was framed in terms of accuracy, and so will be referred to as accuracy ratings in the remainder of this paper. The different groups of listeners rated vowel accuracy or whole word accuracy. Vowel ratings were obtained to evaluate whether or not vowels were hyperarticulated in clear speech. Whole word ratings were obtained in case clear speech effects are realized over larger domains than the segment in preschool children's speech.

2.2.1. Procedure

The rating task was completed by one listener at a time in a sound attenuated booth. A listener received a list of all the target words in a predetermined randomized order. Each word was listed with a 9 point number scale—a scale presumably sensitive enough to capture small between age group and speaking style differences (see, e.g., [10]). The scale was anchored at the top of each page with the words *least accurate* and *most accurate* appearing above numbers 1 and 9, respectively. *Most accurate* was defined either as (1) the vowel in a clearly articulated adult version of the target word, or as (2) the clearly articulated adult version of the target word. The particular definition corresponded to whether listeners were rating the vowel or whole word. *Least accurate* was not defined. Instead, listeners were encouraged to use as much of the scale as possible when making their judgments.

Words were presented auditorily to listeners in one of several predetermined random orders. Individual listener ratings were averaged across each group. Mean accuracy ratings were used in the analyses presented below.

3. RESULTS AND DISCUSSION

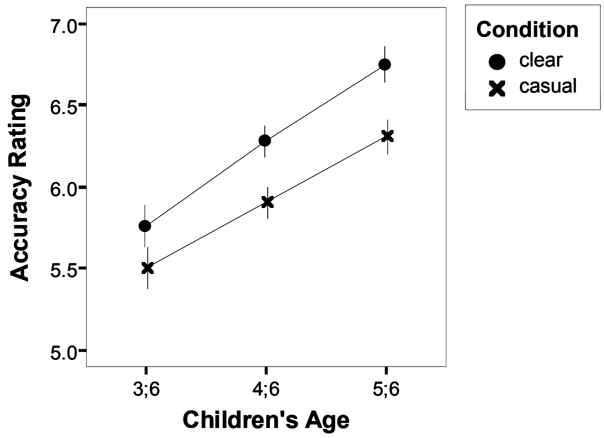
The results indicated that a clear speech strategy develops over time, and that the early strategy may be different in children than in adults. In particular, vowel accuracy ratings and vowel acoustics indicated small style-dependent differences in 4 and 5 year old speech, suggesting that children may not aim for paradigmatic distinctiveness. The word accuracy ratings provided some support for this suggestion in that they showed much larger style-dependent differences in 5 year old speech.

3.1. Perceived Vowel Accuracy

A mixed-model ANOVA assessed the repeated measures effect of speech style on perceived vowel accuracy. Age, vowel front-backness, and vowel height were between group factors. The analysis indicated an effect of condition on mean accuracy ratings [$F(1, 256) = 11.67$, $p < .01$] that interacted with vowel front-backness and height [$F(1, 256) = 4.59$, $p < .05$]. Front and central vowels received higher ratings than the back vowels and the high-front and mid-central vowels received higher ratings than all other vowels.

Age also had a significant effect on the perceived accuracy of vowels [$F(2, 256) = 20.85$, $p < .01$]: listeners gave higher accuracy ratings to vowels produced by older children than to those produced by younger children. Although the interaction between age and speaking style was not significant, post hoc tests ($\alpha = .017$) indicated that listeners rated 4 and 5 year olds' clear speech vowels as more accurate than their casual speech vowels, but listeners' mean accuracy ratings on 3 year olds' vowels were the same regardless of speech style condition. These effects are evident in Figure 1.

Figure 1: Mean vowel accuracy ratings on a nine point scale displayed by speech style condition and the children's age.



3.2. The Vowel Space

The next set of analyses explored whether or not the effect of speech style on vowel accuracy could be attributed to hyperarticulation. A mixed-model ANOVA on F1 measures indicated simple effects of the between-group factors age and vowel height [age, $F(2, 269) = 4.54$, $p < .05$; vowel height, $F(2, 269) = 171.33$, $p < .01$] as well as a nearly significant effect of the repeated-measures factor, speech style [$F(1, 269) = 3.65$, $p = .06$]. More importantly, the analysis indicated a significant interaction between age, vowel height, and speech style [$F(4, 269) = 2.49$, $p < .05$] that was consistent with the rating results. This interaction is shown in Figure 2.

Figure 2: F1 midpoint for high, mid, and low vowels produced by 3, 4, and 5 year old children the clear and casual speech conditions.

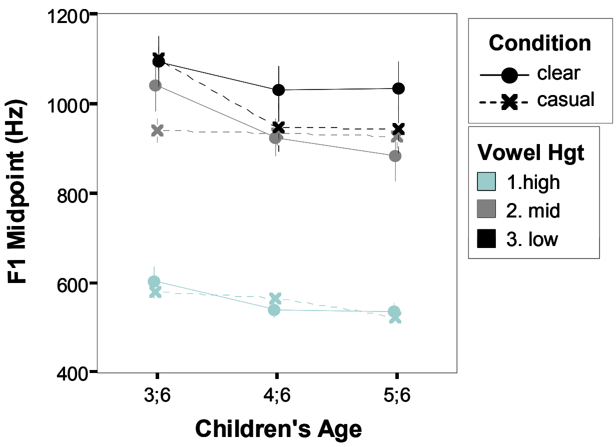


Figure 2 shows that 4 and 5 year olds' high, mid, and low vowels were more distinct from one another in the clear speech condition than in the casual speech condition. This result suggests that young children control the open-close dimension of speech production and can exploit this control to achieve increased speech clarity.

Analysis of the F2 measures provided less compelling results in that they provided little explanation for the differences in perceived vowel accuracy. A mixed-model ANOVA indicated that F2 was somewhat higher in the clear speech condition compared to the casual speech condition [$F(1, 260) = 3.92$, $p < .05$], but this effect did not interact with vowel front-backness, height, or age. In other words, clear speech vowels were somewhat fronted compared to casual speech vowels, but the horizontal vowel space was not expanded.

The other vowel measures did not explain the style-dependent differences in perceived vowel accuracy either. A mixed-model ANOVA exploring the effect of speech style on the distance (in Hz) of F2 onset transitions indicated nearly significant effects of style and of age × style on the onset transitions, but not in the expected direction. The distance from F2 onset to F2 midpoint tended to be greater in the casual speech condition than in the clear speech condition [$F(1, 260) = 3.5$, $p = .06$]. Further, only 3 and 5 year olds showed this difference, 4 year olds did not. Similarly, a mixed-model ANOVA on vowel duration revealed a significant effect of speech style [$F(1, 260) = 6.52$, $p < .05$], but in the opposite direction one might expect based on the adult literature (see, e.g., [8]). Vowel durations were longer in the casual speech condition than in the clear speech condition. This effect of style did not interact with any of the other factors in the analysis. Finally, there were no significant effects of speech style in an analysis of F2-F1 and no significant interaction of speech style with either age, vowel front-backness, or vowel height.

In summary, the acoustic measures on vowels suggest that preschool children may use the open-close dimension in speech to effect a style change, but there is little evidence to suggest that the children aim for extreme acoustic vowel targets in clear speech.

3.3. Perceived Word Accuracy

If preschool children are not aiming for phonemic distinctiveness in clear speech, how do they achieve the style differences evident in the vowel accuracy ratings? One possibility is that children aim for an adult-like word target in clear speech. In so doing, children more clearly enunciate the whole word, which has secondary effects on the perceived accuracy of individual phonemes. The accuracy ratings on whole words provide preliminary support for this idea in that the development of style shifting is better delineated by these ratings compared with the vowel accuracy ratings.

A mixed-model ANOVA was used to assess the repeated measures effect of speech style on perceived word accuracy. Age and syllable structure (simple vs. complex) were treated as between group factors. The analysis indicated a strong effect of condition on mean accuracy ratings [$F(1, 268) = 15.42$, $p < .01$] that interacted with age [$F(2, 268) = 3.24$, $p < .05$], but not with syllable structure. The simple effect of age was significant [$F(2, 268) = 14.29$, $p < .01$], but the effect of syllable structure was not [$F(1, 268) = .03$, NS].

Figure 3: Mean word accuracy ratings on a nine point scale for monosyllabic words produced by preschool children in different speech style conditions.

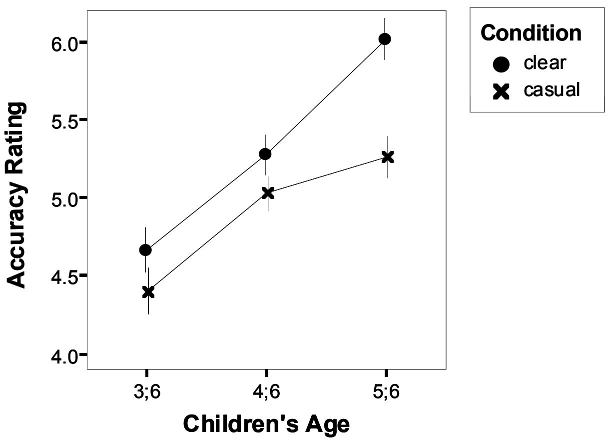


Figure 3 displays the interaction between speech style and age on word accuracy ratings. Post hoc tests ($\alpha = .017$) confirm what is suggested by the figure; namely, that (a) there was no style-dependent difference in ratings of 3 year olds' words, (b) there was an incipient style-dependent difference in ratings of 4 year olds' words, and (c) a large and statistically significant style-dependent difference in ratings of 5 year olds' words. This pattern of results on word accuracy can be contrasted with those obtained on vowel accuracy (Figure 1). Whereas differences in vowel accuracy ratings increased nearly monotonically with age and accuracy differences between speech styles was small (an average of 0.36 points), word accuracy ratings on 5 year old speech show a discontinuity between age groups and speech styles. Five year old casual speech words were rated the same as 4 year old productions, but 5 year old clear speech words were perceived to be more similar to an adult-like clear speech target. This large effect of style on word accuracy ratings suggests that 5 year olds have adopted a clear speech strategy—albeit one that enhances the whole word rather than the phonemic contrasts present in the word.

4. CONCLUSION

Style-dependent differences are more evident at the word-level in 5 year old speech than at the phonemic level. This result may suggest a clear speech strategy aimed at syntagmatic clarity rather than paradigmatic distinctiveness. Age-dependent differences in accuracy ratings suggest that, whatever the strategy, distinct clear and casual speech styles develop slowly over time.

5. REFERENCES

[1] Beckman, M.E., Edwards, J. 2000. The ontogeny of phonological categories and the primacy of lexical learning in linguistic development. *Child Dev.*, 71, 240-249.

[2] Bradlow, A. R., Torretta, G.M., Pisoni, D.B. 1996. Intelligibility of normal speech I: global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication* 20, 255-272.

[3] Johnson, K., Flemming, E., Wright, R. 1993. The hyperspace effect: phonetic targets are hyperarticulated. *Language* 69, 505-528.

[4] Lee, S., Potamianos, A., Narayanan, S. 1999. Acoustics of children's speech: developmental changes of temporal and spectral parameters. *J. Acoust. Soc. Am.* 105, 1455-1468.

[5] Lindblom, B. 1990. Explaining phonetic variation: a sketch of the H&H theory. In: Hardcastle, W., Marchal, A. (eds.), *Speech Production and Speech Modeling*. Dordrecht: Kluwer.

[6] Lindblom, B. 1992. Phonological units as adaptive emergents of lexical development. In: Ferguson, C., Menn, L., Stoel-Gammon, C. (eds), *Phonological development: Models, research, implications*. Timonium, MD: York Press.

[7] McCune, L., Vihman, M.M. 2001. Early phonetic and lexical development: a productivity approach. *J. Speech, Lang. Hear. Res.* 44, 670-684.

[8] Moon, S-J., Lindblom, B. 1994. Interaction between duration, context, and speaking style in English stressed vowels. *J. Acoust. Soc. Am.*, 96, 40-55.

[9] Picheny, M.A., Durlach, N.I., Braida, L.D. 1986. Speaking clearly for the hard of hearing II: acoustic characteristics of clear and conversational speech. *J. Speech Hear. Res.* 29, 434-446.

[10] Southwood, M.H., Flege, J.E., 1999. Scaling foreign accent: direct magnitude estimation versus interval scaling. *Clinical Linguistics & Phonetics* 13, 335-349.