ACQUISITION OF VOWEL DURATION: A COMPARISON OF SWEDISH AND ENGLISH

Carol Stoel-Gammon, Eugene H. Buder, and Margaret M. Kehoe
Speech and Hearing Sciences, University of Washington, Seattle, Washington, USA

ABSTRACT

This study compares durations of the high front unrounded vowels produced by 30-month-old subjects from two language communities: Swedish and American English. The findings indicate that intrinsic and extrinsic factors have different influences on vowel length in the two languages. In productions of children acquiring English, the extrinsic factor of voicing of the final consonant manifests strong effects on the vowel duration, but the intrinsic effects of the tense-lax distinction are absent. For Swedish children, the intrinsic effects are very strong, but there is no effect of consonant voicing.

INTRODUCTION

In acquiring the phonology of their mother tongue, children must learn not only the articulatory gestures needed for the correct production of the consonants and vowels of their language, but also the rhythm and timing associated with sounds, words and phrase.. Acquiring adult-like durational patterns is difficult in part these patterns vary considerably across speaker and context. The length of speech sounds, particularly vowels, is influenced by a number of intrinsic and extrinsic factors, including vowel quality; rate of speech; word and phrasal stress; position within the phrase; and consonantal context. Consequently, the child's is faced with more than simply learning a fixed duration for each segment or word.

In conjunction with researchers from Stockholm University, we have collected and are analyzing a large set of data from infants and toddlers aged 6-30 months. The goal of this joint research project is to examine the development of language-specific speech patterns in the two languages at both the segmental and suprasegmental levels. (See for example the comparison of acquisition of alveolar vs. dental /t/s in 30-month-old subjects from Seattle, USA and Stockholm, Sweden [1].) The investigation reported here focuses on aspects of the acquisition of vowel duration in the two languages.

Vowel Systems

Comparisons between the vowel systems of Swedish and American English are particularly interesting for crosslinguistic studies. The Swedish vowel system is typologically large, with up to 18 distinct vowels in contrast to the American English system which contains 12 vowel phonemes (excluding diphthongs). The Swedish system is typically described as being composed of nine short-long pairs, although quality differences in each pair do exist. [2,3]. The phonemic system of American English, by comparison, is generally described as having vowels that are distinguished primarily by differences in quality.

Previous investigations of vowel duration in the two languages have highlighted the presence of both intrinsic and extrinsic differences in duration. **Intrinsic** differences are readily apparent in tense-lax (or long-short) vowel pairs. For both languages, the vowel referred to as "tense" is longer than its "lax" counterpart in contexts in which other phonetic features are held constant. Thus, for example, in English the lax vowel /I/ of the word "bit" is shorter than the tense vowel /i/ of "beat". (Note: The IPA symbol [I] for the lax vowel used here corresponds to the Swedish Technical Alphabet short vowel i2.).

Extrinsic differences in vowel duration are conditioned primarily by phonetic features of the consonant which follows the vowel; if all other parameters are held constant, vowels tend to be shorter when they precede voiceless consonants than in other environments. Thus, in English, the /i/ of "beat" is shorter than the /i/ of "bead"). The finding that vowels are shorter before voiceless obstruents than before voiced obstruents and sonorants has been documented in a wide range of languages, leading some researchers to suggest that this particular durational difference is universal and physiologically determined (e.g., [3, 5].

However, Keating [6] points out that in a small number of languages, including Polish, Czech and Saudi Arabic, there is no interaction between vowel length and consonant voicing. Consequently, she argues that context-sensitive patterns of vowel duration represent rule-governed behaviors that are language specific and must be learned.

In the case at hand, American English and Swedish are among the large group of languages that follow the general pattern of longer vowels before voiced consonants. At the same time, as shown below, the two languages differ substantially in the magnitude of the effect on vowel duration.

Adult Data: Intrinsic Patterns

Detailed work by Elert [3] has shown that the intrinsic durational difference between lax and tense vowels in adult Swedish are substantial. The ratio of lax-to-tense vowels, averaged across all pairs, is .65, meaning that the lax vowels are about 2/3 as long as their tense counterparts. For American English, the durational differences are somewhat less marked with a lax-to-tense ratio around .71 according to House [4]. As noted above, in most phonological analyses of American English, intrinsic durational differences are considered to be secondary cues for distinguishing tense-lax pairs such as /i/ vs. /I/ or /u/ vs. /U/; the primary distinction stems from formant (i.e., quality) differences in such pairs.

Adult Data: Extrinsic Patterns

Data from existing studies of extrinsic vowel quantity effects indicate that voicing of the following consonant exerts a strong influence on vowel duration in American English. Vowels preceding voiced consonants are nearly twice as long as the same vowels before voiceless consonants. House [4] (see also Hoard [7]) reports that, on average, the ratio of vowels preceding voiceless compared with voiced consonants in monosyllabic words is .51.

In Swedish, the extrinsic effects of consonant voicing are much smaller, as might be expected in a language with a phonemic vowel length contrast. (If vowel durations were to shift substantially as a result of voicing of the following consonant, intrinsically long vowels might be perceived as their short/lax counterparts and vice-versa.) For adult Swedish, Elert [3] reports an average voiceless-to-voiced ratio of .97, indicating little influence of the voicing status of the following consonant.

If we compare the two languages, then, it is clear that intrinsic differences are strong in adult Swedish, and are present, but less marked in adult American English. Extrinsic differences in vowel length, by comparison, are very strong in American English, but are minimal in Swedish.

Child Data

To date, research on the acquisition of vowel length by children acquiring American English has provided inconsistent findings. Naeser [8] reported that the intrinsic and extrinsic vowel length ratios of her 22-month-old subjects were similar to those of the adults in her study. The children's intrinsic ratio was .74 (compared with an adult ratio of .71); the children's extrinsic ratio was .50 (cf. an adult ratio of .59). A year later, however, the children's ratios were less adult-like and differed minimally from one another:: The average intrinsic ratio was .68 and the average extrinsic ratio was .62. (The ratios given here were calculated on the basis of raw data presented in Naeser's report.)

In a subsequent study of vowel duration patterns in American children, aged 26 months, Greenlee [9] found that intrinsic and extrinsic duration patterns were approximately the same: .68 for the intrinsic ratio and .66 for the extrinsic ratio.

In sum, the data on acquisition of vowel duration by American children are unclear. At 22 months, the children in Naeser's study exhibited extrinsic differences that were more marked than

intrinsic differences, a pattern that conforms to adult speech. The findings for older children, however, indicate equal levels of shortening in both contexts, unlike adults.

We know of no data on the acquisition of vowel duration in children learning Swedish. Given the marked differences in the intrinsic and extrinsic durational patterns of Swedish and American English vowels, we might expect that adult-based patterns would emerge early in child speech. On the other hand, it is possible that all children will follow a similar developmental course, with intrinsic and extrinsic differences acquired in a fixed order.

Purpose

The purpose of this study was to examine the acquisition of intrinsic and extrinsic vowel length patterns in Swedish and American English by comparing productions of the vowels /i/ and /I/ in the speech of young children acquiring these two languages. The vowel pair /i/ - /I/ was selected for a number of reasons:

(1) In Swedish and English, these vowels form a tense-lax (or long-short) pair, making them a comparable match across the two languages;

(2) These vowels occur frequently in words that are part of young children's vocabulary and thus spontaneous productions could be elicited;

(3) Previous research suggests that these vowels are acquired relatively early in child speech [8,10].

Predictions

Given the findings of previous investigations of adult speech, we formed a set of predictions regarding the develop^mental patterns of intrinsic and extrinsic vowel duration in the speech of Swedish and American children:

(1) The vowels of *Swedish* children would be influenced more by intrinsic than extrinsic lengthening, as in the adult model.

(2) Extrinsic vowel length differences would be greater in the speech of *Swedish* children than in the models, because the children would have to learn to overcome the apparently universal pattern of lengthened vowels preceding voiced consonants

(3) Both intrinsic and extrinsic duration differences would be present in the productions of *American* subjects.

(4) Extrinsic differences would be greater than intrinsic differences for the *American* children, as they are in the adult model.

METHOD

Data Collection

Subjects included 18 children aged 30 months: nine acquiring Swedish in Stockholm; nine acquiring American English in Seattle. At both sites, the subjects participated in semi-structured tasks during which words with /i/ and /I/ followed by voiced and voiceless consonants were elicited. Care was taken to elicit monosyllabic forms with obstruent onsets and offsets to aid subsequent durational analyses. In addition, an effort was made to elicit phonetically similar words (including some nonsense forms) in the two languages.

All speech samples were recorded in sound-treated rooms using Lavalier microphones placed in a soft vest worn by the subjects. The microphone was linked to an FM wireless system. At both sites, the speech signals were recorded on Panasonic VHS videocassette recorders using High-Definition audio tracks.

Database

The data base for the present study consisted of productions of isolated or phrase-final words of the shape CVC in which the vowel was /i/ or /I/ and the inital and final consonants were obstruents. Because Swedish has few monosyllables in which lax /I/ is followed by a voiced consonant, few exemplars of this form were obtained; consequently this category was not included in the durational analyses for either language.

The children's productions were divided into three groups based on durational patterns:

Group 1 consisted of words with lax /I/ followed by a voiceless consonant;

Group 2 consisted of words with /i/ followed by a voiceless obstruent;

Group 3 consisted of words with /i/ followed by a voiced consonant.

Analyses of *intrinsic* durations are based on comparisons of Group 1 vowels with Group 2 vowels. Analyses of *extrinsic* duration involve comparisons between the vowels in Groups 2 and 3.

Acoustic Measures

All words were digitized and vowel durations were measured using the spectrographic display produced by the Computerized Speech Lab (Model 3400, Version 4.0 Kay Elemetrics). After eliminating tokens that were not acoustically analyzable because of poor voice quality or noise overlay, the vowel durations of 206 monosyllables were measured using the following criteria:

(1) Vowel onset was indicated by released vowel energy showing clear periodicity and energy in the first three formants.

(2) Vowel offset was indicated by the evidence of oral closure (i.e., a sudden reduction in waveform envelope and a loss of clear formant energy).

The first three formants of all vowel tokens were also measured using procedures developed by Buder and Stoel-Gammon [11].

RESULTS

Duration Measures

Average durations for the American and Swedish subjects for the vowels of interest are as follows.

Group 1 vowels: The mean duration of /I/ followed by a voiceless consonant was 202 ms for the American subjects and 161 ms for the Swedes.

Group 2 vowels: The mean duration of /i/ followed by a voiceless consonant was 191 ms for the American subjects and 326 ms for the Swedes.

Group 3 vowels: The mean duration of /i/ followed by a voiced consonant was 329 ms for the American subjects and 295 for the Swedes.

Intrinsic and Extrinsic Ratios

Figure 1 (next page) presents individual and group data on intrinsic and extrinsic vowel duration ratios. The top half of the figure shows the findings for the American subjects; the bottom half provides the same data for the Swedish subejcts.

Sections (a) present the individual ratios for extrinsic (Ext) and intrinsic (Int) durations for the nine subjects in each group;

Sections (b) present averaged group data for the individual ratios with error bars showing the standard error of the mean;

Sections (c) show formant measures (in mels) for the two vowels in question. /i/ is indicated by triangles and /I/ by circles.

The *intrinsic* ratios presented in Figure 1 are based on comparisons of /I/ and /i/ before a voiceless obstruent. The average intrinsic ratio for the American children was 1.06 [see Figure 1, Section (b), top half]. The average intrinsic ratio for the Swedish children was .49 [see Figure 1, Section (b), bottom half].

Average *extrinsic* ratios (based on subjects' averages of /i/ before voiceless and voiced obstruents) for the two groups were .58 for the American children and 1.10 for the Swedes [see Figure 1, Sections (b)].

Statistical Measures

Statistical analyses using the Wilcoxon rank sum test for dependent groups indicated that the extrinsic ratio in American English was significantly smaller than the intrinsic ratio ($p < .05$) [see Figure 1, Section (b), top half]. In Swedish, the intrinsic ratio was significantly smaller than the extrinsic ratio ($p < .05$) [see Figure 1, Section (b), bottom half].

Examination of the individual data in Figure 1, Sections (a) reveals that among the American subjects, two children weakly violated the group pattern by exhibiting intrinsic and extrinsic ratios that were essentially equivalent. Among the Swedish children, one subject strongly violated the pattern used by others in the group.

Quantity vs. Quality

The finding that there was little difference in the durational values for /I/ and /i/ in the productions of American children (with a lax-to-tense ratio of 1.06) raises the possibility that these vowels were indistinguishable in their speech. Acoustic analyses of formant structure indicated, however, that the children's /I/s and /i/s were clearly

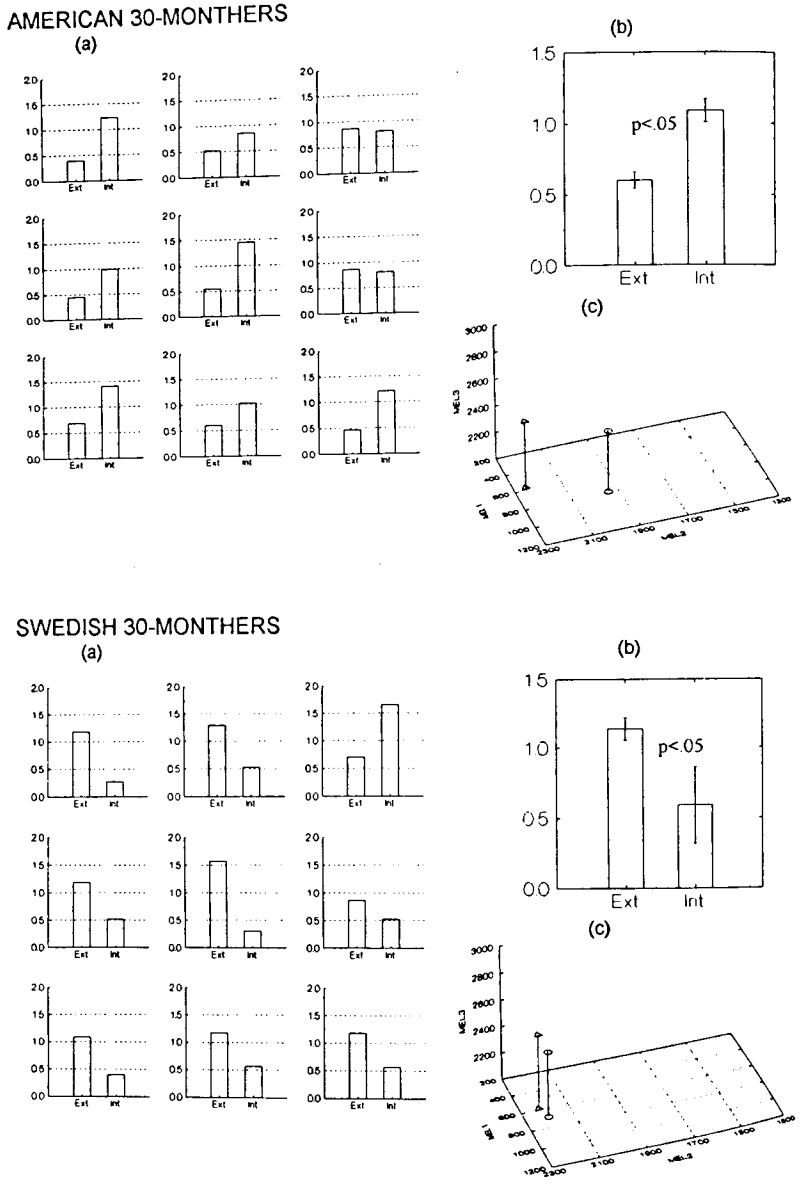

Figure 1. Comparison of quantity and quality measures of /i/ and /I/ from CVC words produced by 9 American (top panel) subjects , aged 30 months, and 9 Swedish subjects (bottom panel) aged 30 months. Sections (a) show individual data; sections (b) and (c) show group averages. See text for a full explanation.

different in terms of quality. As shown in panel (c) in the upper half of Figure 1, the /i/ tokens produced by the American subjects were characterized by a lower F1 and a higher F2 than /I/, as is the case in adult speech.

In contrast, the /I/ and /i/ productions of the Swedish children exhibited little difference in quality, as shown in the lower half of Figure 1, Section (c). Rather, the vowels were distinguished by differences in quantity -- on average, /i/ was more than twice as long as /I/.

DISCUSSION

Comparisons with Adult Data

If we compare the children's ratios with those of adults, we see that the patterns are roughly similar, though certainly not identical. For adult Swedish, Elert [3] reported an ratio of .68 for /I/ and /i:; the children's ratio from this study was .49. The extrinsic ratio for /i/ before voiceless and voiced consonants in adult speech was .99 according to Elert, compared with 1.10 for the Swedish children.

For American English, House [4] reports an adult intrinsic ratio for /I/ vs /i/ in monosyllables of .74, compared with the children's ratio of 1.06. The adult extrinsic ratio for /i/ before voiceless and voiced consonants was .49 for adults and .58 for children in this study.

Predictions

Four predictions regarding vowel length patterns in the speech of young Swedish and American children were made at the outset of the study. Some were supported; other were not. The predictions are repeated below and compared with the findings.

(1) *The vowels of Swedish children would be influenced more by intrinsic than extrinsic lengthening, as in the adult model.*

This prediction was strongly supported. The children's productions exhibited a very strong influence of intrinsic shortening; in fact, the influence was much stronger in the children's productions than in the adults from Elert's study.

(2) *Extrinsic vowel length differences would be greater in the speech of Swedish children than in the adult models, because the children would have to learn to overcome the apparently universal pattern of lengthened vowels preceding voiced consonants.*

This prediction was not supported. On average, the Swedish children produced longer vowels before voiceless consonants than before voiced ones, in spite of the fact that most languages of the world conform to the predicted pattern. This finding raises questions about the physiological underpinnings of vowel lengthening preceding voiced consonants.

(3) *Both intrinsic and extrinsic duration differences would be present, in the productions of American subjects.*

This prediction was not supported. In the speech of the American children, the lax vowel /I/ was, on average, slightly longer than the tense vowel /i/, contrary to the adult pattern. As noted above, the primary distinction between /I/ and /i/ was in quality rather than quantity.

(4) *Extrinsic differences would be greater than intrinsic differences for the American children, as they are in the adult model.*

This prediction was supported.

Future Research

The findings from this study are intriguing and suggest several avenues of future research:

(1) Studies of younger and older children are needed to trace the development of language-specific durational patterns from emergence to mastery.

(2) Investigations of other vowels are needed to determine if the findings for the pair /I/ - /i/ hold true across the vowel system. Of particular interest in this regard would be an examination of other Swedish pairs that differ to greater extent in quality. It may be that durational differences are acquired later in such pairs since phonmic contrast is marked by quality az well as quantity.

(3) Information on individual patterns of development would be useful in comparing the roles of quality and quantity in the acquisition of the vowel system. For example, in this study, the durations of one Swedish child differed dramatically from those of her peers. It is possible that she was using vowel

quality rather than quantity to create vowel contrasts.

(4) Data on the acquisition of vowel durations in disyllabic words are needed for a full understanding of the influence of phonetic context and word length on intrinsic and extrinsic patterns.

(5) Finally,investigations of adult-child interactions are needed to determine the durational patterns of vowels in child-directed speech.

ACKNOWLEDGMENTS

This research was supported in part by a grants from NIH-NIDCD, the Royalty Research Fund and the Virginia Merrill Bloedel Hearing Research Center at the University of Washington. Data collection was carried out in collaboration with Karen Williams and Olle Engstrand of the Institute of Linguistics, University of Stockholm.

REFERENCES

[1] Stoel-Gammon, C., Williams, K., & Buder, E.H. (1994) Crosslanguage differences in phonological acquisition: Swedish and American /l/. *Phonetica,* vol. 51, pp. 146-158.

[2] Fant, G. (1973). *Speech sounds and features.* Cambridge, MA: MIT Press.

[3] Elert, C-C. (1964) *Phonological studies of quantity in Swedish.* Uppsala: Almqvisst & Wiksells.

[4] House, A.S. (1961). On vowel duration in English. *The Journal of the Acoustical Society of America,* vol. 33, pp. 1174-1178.

[5] Chen, M. (1970). Vowel length variation as a function of the voicing of the consonant environment. *Phonetica,* vol. 22, pp. 129-159.

[6] Keating, P.A. (1984). Phonetic and phonological representation of stop consonant voicing. In V. Fromkin (Ed.), *Phonetic linguistics.* New York: Academic Press.

[7] Hoard, J. E. (1966). *Contrastive analysis of English and Swedish phonology.* Unpublished master's thesis, University of Washington, Seattle, WA.

[8] Naeser, M. A. (1970). *The American child's acquisition of differential vowel duration.* (Tech. Rep. No. 144). Madison, WI: University of Wisconsin, Center for Cognitive Learning.

[9] Greenlee, M. (1977). *Another look at the acquisition of differential vowel duration in American English.* Unpublished doctoral dissertation, University of California, Berkeley.

[10] Otomo, K. & Stoel-Gammon, C. (1992). The acquisition of unrounded vowels in English. *Journal of Speech and Hearing Research,* vol. 35, pp. 604-616.

[11] Buder, E., & Stoel-Gammon, C. (1993). Obtaining valid and reliable acoustic measures of children's vowel productions. Paper presented at the American Speech-Language-Hearing Association, Anaheim, CA, November.