

Acquisition of Vowel Duration in Children Speaking American English

Eon-Suk Ko

Department of Cognitive and Linguistic Sciences, Brown University, U.S.A.

esko@alumni.upenn.edu

Abstract

This study is an acoustic investigation of the acquisition of vowel duration in children speaking American English. The primary goal was to find out when and how children begin to produce different vowel durations as a function of postvocalic voicing. A total of 803 longitudinal data extracted from the Providence Corpus were analyzed. The age range covered by the data was from 0;11 to 4;0. The findings are summarized as follows: (1) Children control the vowel duration conditioned by voicing before the age of 2. (2) They also make the durational distinction between the tense and lax vowels before the age of 2. (3) There is no developmental trend in the acquisition of the vowel duration conditioned by postvocalic voicing. The results suggest that children thoroughly learn the phonetic implementation of temporal parameter from the very early stage of speech production to such an extent as to make it appear as an automatic process.

Index Terms: vowel duration, voicing, tense/lax, corpus, acquisition, child(ren)

1. Introduction

The fact that English vowels are longer before voiced than voiceless consonants (henceforth referred to as the *vowel-length effect* or VLE, following [1]) has been known for a long time ([2-5]). This tendency is found in many other languages ([3]), leading some researchers to consider this effect as a universal phonetic process. Exceptions, however, have been reported from Spanish ([2]), Polish, Czech, and Saudi Arabic ([6]), suggesting that it may be a language-specific phonetic pattern that has to be learned.

Each of the two positions on the VLE makes a different prediction on the acquisition of vowel duration. If VLE is a universal or an automatic process, we would expect children to exert the VLE from the very beginning of coda production. Should there be, however, a stage where children undergo a trial and error period in acquiring the VLE, it would support the view that it is a learned process. Importantly, however, the absence of the learning curve does not necessarily preclude the possibility of the VLE being a controlled process because children could still have to learn it before they actually utter it. I will make a further discussion on this point in section 4.

This study examines the acquisition of the VLE in children's speech to address the question of automatic vs. controlled process on the VLE. The immediate goal of the acoustic analysis is to find out the earliest age when children demonstrate the VLE, and whether there is any developmental pattern in its acquisition. While the main concern of the present study is with the VLE, it also analyzes the pattern of vowel duration in the tense and lax vowels as it may have relevance in answering the question of whether the VLE is a controlled or an automatic process.

Previous research has been made on the development of the production of vowel duration as a function of postvocalic voicing ([7-9]). Highlights of the findings made in previous

studies include the following: (1) The VLE is demonstrated as early as 24 months ([8]). (2) On the other hand, the duration difference between tense and lax vowels was not shown until 30 months [9]. (3) Vowel duration decreases with the increase of age in front of a voiced consonant. The opposite trend was not found before a voiceless consonant ([7]).

The current study is set apart from the previous ones in the following regards: (1) **DATA RANGE:** The range of the age covered in the current study extends to a much earlier age of about 11 months than in previous studies. (2) **DATA COLLECTION:** The data analyzed for the current study are extracted from a corpus database rather than recorded in a laboratory. (2) **RESEARCH DESIGN:** The study is based on longitudinal data (ages 0;11 to 4;00) versus cross-sectional data used in [7] (children of ages 3 and 6, and adults), [8] (around 24 months), and [9] (around 24 and 30 months).

2. Method

2.1. Providence Corpus

The data used in this study were extracted from a subset of recordings in the Providence Corpus ([10]), which contains longitudinal spontaneous child-adult speech interactions of six children from southern New England. Digital audio/video recording was made for approximately one hour every two weeks, commencing with the onset of children's first words. The speech was then transcribed according to the CHAT conventions ([11]) to be included in the CHILDES database. Analyses were made of the speech data from 4 children, who are all monolingual speakers of Standard American English.

2.2. Data

Target words for analysis, presented in Table 1, were selected from the frequency list of the words in each child's speech. Only a subset of the available tokens were analyzed for several reasons including overlapping speech, segmentation difficulties due to noise or unclear articulation, and unusual behaviors such as vocal play. Minimal or near-minimal pairs of the form (C)VC that occur in isolation or in final position of a prosodic phrase were selected for measuring the vowel duration as a function of voicing and tense/lax distinction. In addition, several individual words that do not occur in minimal pairs but are high in frequency were selected to complement the observation of longitudinal vowel duration developmental patterns.

2.3. Acoustic analysis

Segmentation of a vowel was made based on a wide-band spectrogram in consultation with the waveforms using Praat. In generating the spectrogram, the time constant was set at 0.003 sec. given the high pitch of children's voice. The beginning of the vowel was marked at the initiation of the formant structure indicated by the shift of overall energy. Devoiced vowels or aspiration due to the delayed voicing of

Table 1. (Near) Minimal pair data analyzed for the vowel duration differences and results of the t-tests.

	(near) minimal pair data			t-test (total)			t-test (under 2;00)		
	child	word 1 (n=total, under 2;00)	word 2 (n=total, under 2;00)	t	df	p	t	df	p
Vowel Length Effect	Lily	back (n= 62, 11)	bag (n= 25, 8)	-5.0974	58	<0.001	-5.0177	16	<0.001
		duck (n= 63, 39)	bug (n= 18, 7)	-10.09	79	<0.001	-8.277	44	<0.001
		boot (n= 8, 8)	food (n= 27, 6)	-3.5368	33	<0.001	-3.4472	12	<0.005
	Naima	duck (n=33, 16)	bug (n=11, 10)	-3.4675	43	<0.001	-2.2949	24	<0.05
	Ethan	bob, 29(2)	top, 30(11)	-8.1611	45	<0.001	-6.6203	11	<0.001
Violet	back, 33(11)	bag, 25(18)	-4.9902	48	<0.001	-3.7024	27	<0.001	
Tense/Lax Distinction	Lily	sit, 18(4)	seat/feet, 32(15)	4.8649	48	<0.001	2.23	17	<0.05
	Naima	sit, 48(36)	seat, 9(9)	2.1092	42	<0.05	3.5491	38	=0.001
	Ethan	this, 38(5)	piece, 13(9)	3.2709	49	=0.001	1.4013	12	n.s.

the onset consonant were not included in the duration of a vowel. The offset of the vowel was defined by the cessation of the second formant energy or clear overall formant structure. Several irregular instances of the coda such as devoicing, frication or insertion of a vowel after the coda did not exclude the data from measurement as far as the boundaries for segmentation of the vowel were clear as determined by the formant energy or structure.

2.4. Statistical analysis

In many minimal pair groups, the distribution of the vowel duration was close to normal, so I decided to do a parametric test across all data for the coherence of presentation. A two-sample t-test was done for each data group of vowel duration, whose results are summarized in Table 1. The results of the t-test, however, should be taken with caution in some cases because of the potential factors such as the presence of outliers, unequal variance, and the small sample sizes. To supplement the summary of the t-tests, data are also presented in boxplots in section 3.1 and 3.2. A correlation test was made in order to examine the relationship between the vowel duration and age in section 3.3.

3. Results

3.1. Vowel length effect

None of the children showed any obvious developmental stage where they make “mistakes” in exerting the VLE. Rather, children seemed to have no difficulty throughout the examined age range in producing the vowel with varying duration depending on the postvocalic voicing. A case of Lily is illustrated in Figure 1.

The earliest age so far reported for children producing the VLE is 24 months ([8]). It was found, however, that this effect is already present before the age of 2. Figure 2 is a boxplot of vowel tokens from Lily’s speech which were produced before she turned 2. The tendency of a vowel being produced longer before a voiced than a voiceless consonant is demonstrated clearly across different vowel qualities.

Figure 3 presents additional data from Naima, Ethan, and Violet, where the VLE is consistently demonstrated across speakers even before they reach the age of 2.

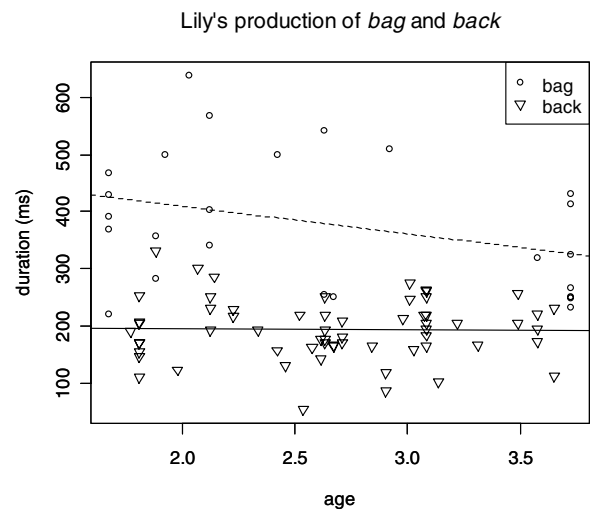


Figure 1: The development of vowel duration before voiced and voiceless consonants in Lily

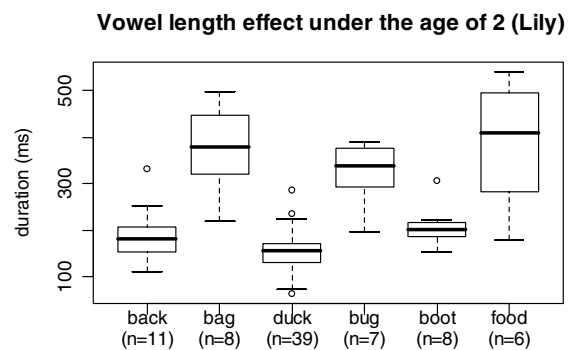


Figure 2: Vowel duration as a function of postvocalic voicing under the age of 2 (Lily)

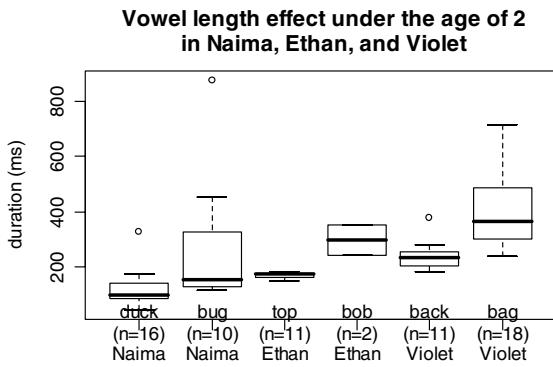


Figure 3: Vowel duration as a function of postvocalic voicing under the age of 2 (Naima, Ethan and Violet)

The earliest age the VLE is encountered in each child somewhat varies. In the case of Naima, the average duration of the vowel in *bug* produced at the age of 1;4 was 256.97 ms. (n=7) whereas that in *duck* produced at the age of 1;6 was 106.45 ms. (n=2). In the case of Violet, the corresponding tendency was found at the age of 1;6 and 1;8. Given that these ages are the temporal points where a suitable minimal pair token was first encountered, it may be that children were already exerting the VLE even before these ages, possibly as soon as they produced a coda.

3.2. Duration in tense and lax vowels

In section 3.1, children’s speech production demonstrated no obvious stages of difficulty in the acquisition of the VLE. This finding seems to make the “automatic” perspective more favorable than the “controlled” one.

In this section, I shift the focus onto the development of vowel duration as a function of the *tense* and *lax* distinction. It is well-known that tense vowels are realized longer than lax vowels, other things being equal. Should we find a developmental stage where children make “mistakes” in producing the vowel duration as a function of the tense and lax distinction, this would support the potential claim that the VLE observed in section 3.1 is truly an automatic process in contrast to the tense/lax distinction which would be a controlled, or a learned process. If, however, we do not find any such developmental stage, explaining the tense/lax effect on vowel duration as a controlled process would require an elaborate argumentation.

I analyzed the duration of tense and lax vowels of (near) minimal pair words occurring in isolation or phrase final position from the speech of Lily, Naima, and Ethan. Figure 4 illustrates the case of Lily’s production of the *seat/feet* and *sit* pair. It seems that Lily has a clear tendency to produce a tense vowel with longer duration than a lax vowel from early on.

Again, in order to find out how early children show the control over the distinction between the tense and lax vowel durations, I analyzed a (near) minimal pair from the speech of Lily, Naima, and Ethan produced before 2;00. Violet was not included here as she did not have any suitable minimal pairs. Figure 5 demonstrates the distribution of these data in a boxplot. We can observe the tendency for a tense vowel to be produced longer than a lax vowel across all speakers.

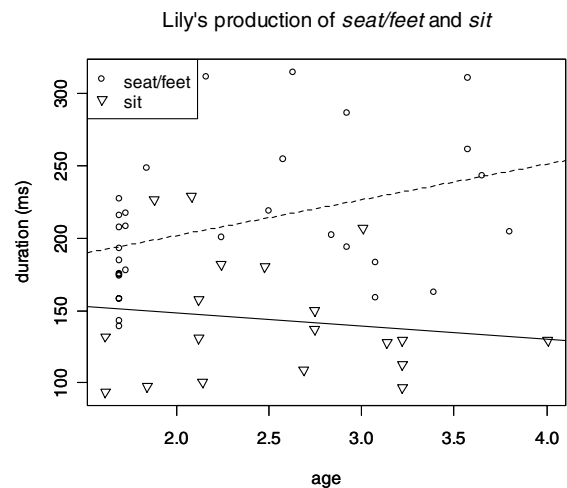


Figure 4: The longitudinal development of tense and lax vowel durations in Lily

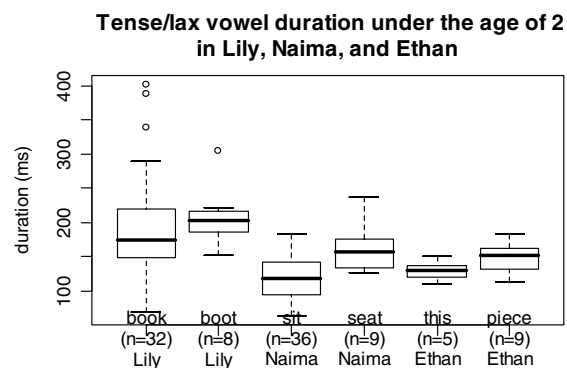


Figure 5: Vowel duration as a function of tense/lax distinction under the age of 2 (Lily, Naima, Ethan)

This result is somewhat at odds with the finding in [9], where children did not show the vowel duration difference for tense/lax vowels until 30 months. The discrepancy could be due to the nature of the data group examined in the two studies, i.e. longitudinal vs. cross-sectional.

3.3. Developmental trend of the vowel length effect

In order to examine whether there is any meaningful trend in the longitudinal development of the VLE, it is ideal to examine data that are continuously present with minimal temporal gaps in observation. Thus I analyzed four most frequent words of the form CVC from the speeches of Naima and Lily, controlling the prosodic position.

Table 2. Correlation statistics for duration and age

	N.	t	df	p	r
cat (Naima)	107	2.7607	105	<0.01	0.26
big (Naima)	71	-1.6835	69	n.s.	-0.20
book (Lily)	103	-3.8423	100	<0.001	-0.36
back (Lily)	62	0.4738	38	n.s.	0.08

In Table 2, only two instances have a significant correlation between duration and age, but their direction is opposite. Thus, the results of the correlation test in Table 2

suggest that the vowel duration does not have any meaningful developmental pattern in children's speech.

4. Discussion

The current study did not find any obvious developmental patterns in the acquisition of the VLE, virtually eliminating this effect from the landmarks in speech development. These results seem to be in favor of the claim that the VLE is an "automatic" rather than "controlled" process. Recall that the previous claim for the VLE being a universal or automatic process was mostly based on the fact that the same tendency is found across many other languages. The results obtained in the current study seem to suggest that the VLE is an automatic process for a different reason, i.e. no learning curve in its acquisition.

As I suggested earlier, however, we cannot rule out the possibility of children learning a phonetic implementation without actually going through a physical trial and error process. In other words, children could have learned the VLE before they first utter the coda. If this were the case, the VLE could still be explained as a controlled process.

The proponents of interpreting the VLE as a universal phonetic process based their arguments on cross-linguistic data (e.g. [3]). As for a mechanism behind the automaticity, they offered production-based explanations which attribute the VLE to the constraints reflecting articulatory contingencies.

Most of the production-based explanations, however, have been convincingly criticized as being vague or based on experiments with erroneous design (see [1] and references therein). Alternatively, the auditory explanations claim that speakers produce the VLE as a strategy to perceptually enhance the voicing feature of the consonants by inducing contrast effects between the vowel duration and the closure-duration of the consonants.

If we accept the auditory view and assume that the VLE is a controlled articulatory strategy to enhance auditory perception, an interesting prediction is made. That is, speakers will have the option of not using the VLE as a cue for voicing. This prediction is consistent with the observation of deviant durational cases such as the Scottish vowel lengthening that occur before voiced fricatives, /r/, and morpheme boundaries in contrast to the environment of voiced consonants in other languages ([12]), and the lack of VLE in Polish voicing distinction ([6], mentioned in [1]), etc.

If the VLE is a controlled process, why is there little robust evidence for it? One explanation for the reason why the VLE appears to be an automatic process in children's speech may be found from the hypothesis of "overlearned" phonetic implementation ([13]). That is, children learn the appropriate control regimes for articulating the vowel in the segmental context so thoroughly that the phonetic implementation process appears to be an automatic process. The absence of developmental trends noted in 3.3 may support this view.

5. Conclusion

The present study discovered that children acquire control over vowel duration in both the postvocalic voicing and the tense/lax specifications well before the age of 2. In addition, no obvious developmental trends were found in the acquisition of the VLE, which suggests that these effects are not learned gradually. Given that the vowel duration is a secondary cue in both voicing and tense/lax features ([14]), it

seems that the VLE is a controlled rather than automatic process in an effort to enhance the auditory cues for these features.

In terms of methodology, the current study show-cases an acoustic study of children's speech using longitudinal data in a speech corpus and is expected to promote similar research in this less investigated area of study.

6. Acknowledgements

I appreciate Katherine Demuth very much for having invited me to work in her lab and allowed me to use the Providence Corpus (funded by NIH Grant #R01MH0922) before its public release, and Jae-Yung Song for giving me a reference. All errors are my own.

7. References

- [1] K. R. Kluender, R. L. Diehl, & B. A. Wright, "Vowel-length differences before voiced and voiceless consonants: an auditory explanation," *JPhon*, vol. 16, pp. 153-169, 1988.
- [2] A. A. House, "On vowel duration in English," *JASA*, vol. 33, pp. 1174-1178, 1961.
- [3] M. Chen, "Vowel length variation as a function of the voicing of the consonant environment," *Phonetica*, vol. 22, pp. 129-159, 1970.
- [4] D. H. Klatt, "Interaction between two factors that influence vowel duration," *JASA*, vol. 54, pp. 1102-1104, 1973.
- [5] L. Lisker, "On "explaining" vowel duration variation," *Glossa*, vol. 8, pp. 223-236, 1974.
- [6] P. A. Keating, "Universal phonetics and the organization of grammars," in *Phonetic linguistics: essays in honour of Peter Ladefoged*, V. A. Fromkin, Ed. Orlando, FL.: Academic Press, 1985, pp. 115-132.
- [7] S. E. Krause, "Developmental use of vowel duration as a cue to postvocalic stop consonant voicing," *Journal of Speech and Hearing Research*, vol. 25, pp. 388-393, 1982.
- [8] C. Stoel-Gammon and E. H. Buder, "Vowel length, post-vocalic voicing and VOT in the speech of two-year olds," *Proceedings of the ICPHS 1999*, pp. 2485-2488, 1999.
- [9] C. Stoel-Gammon and E. H. Buder, "American and Swedish children's acquisition of vowel duration: Effects of vowel identity and final stop voicing," *JASA*, vol. 111, pp. 1854-1864, 2002.
- [10] K. Demuth, J. Culbertson, and J. Alter, "Word-minimality, epenthesis, and coda licensing in the acquisition of English," *Language & Speech*, vol. 49, pp. 137-174, 2006.
- [11] B. MacWhinney, *The CHILDES project: Tools for analyzing talk. 3rd Edition. Vol 2: The database*. Mahwah, NJ: Lawrence Erlbaum Associates, 2000.
- [12] A. McMahon, "Lexical phonology and sound change: the case of the Scottish vowel length rule," *Journal of Linguistics*, vol. 27, pp. 29-53, 1991.
- [13] J. Kingston and R. Diehl, "Phonetic knowledge," *Language*, vol. 70, pp. 419-454, 1994.
- [14] H. Giegerich, *English phonology: an introduction*. Cambridge: C.U.P., 1992.