Development of complex syllable onsets: evidence from durational measurements

Fredrik Karlsson  
Department of Philosophy and Linguistics, Umeå University

Abstract

This paper investigates the emergence of complex syllable onsets through measurements of their duration compared to the duration of corresponding singleton syllable onsets. The results show that there is a difference in the progression in syllable onset duration that is dependent on subject age and onset structure.

Introduction

It has been reported in the literature that young children produce different target words that are perceived by adults as identical (e.g. Spencer 1988, McLeod et al. 1996, Czigler et al. 2000). Yet it is unclear what mechanism lies behind these productions.

The observation that child speech contains more durational variation than adult speech (Kent 1979) has resulted in many viewing this solely as an effect of the maturing speech production system rather than as an effect of a coupling of internal and a maturing speech production system.

Braine (1976) and Macken (1980), among others, have questioned whether when investigating the early stages of child speech development it can be assumed that children have an adult-like representation of the produced word. Spencer (1988) extended this idea when he posited a model of the speech development process that proposes a perceptual filter which constructs an input representation in the child. In production, this input representation is encoded through realization processes as a 'output underlying representation'. Due to the structure of the model, it allows for the possibility of differences between the child's internal representation and the 'output underlying representation'. These differences may be due to interfering processes in the perceptual filter and, at the production end, in the encoding phase.

The two types of models of speech production described above, the models proposing an underlying representation that are identical to the adult output form and the one proposed by Spencer (1988) allowing for a influencing perceptual filters, generate different hypothesis concerning the acoustic properties of original and processed word pairs.

According to the first group of models, the differences between the output form of the child and adult representation is necessarily caused by a not-yet fully-developed production system. In this scenario, two different internal representations might be expected to give rise to acoustic differences in the output pairs. Furthermore, an acoustic feature is expected to be present in the appropriate contexts once the artikulatory proficiency has been developed.

In contrast, according to the Spencer model, homonyms might also be caused by perceptual processes in the child, causing identical internal representations of the two target words. In this case, there is no a priori reason for there to be significant differences between the output forms of the two words since they would have equal opportunity of being affected by production processes. These groups of models allow for acoustic features to emerge at different stages in development according to the level of specification of the child’s underlying form of the target word.

Czigler et al. (2000) compared the syllable onset durations in target words with initial s+plosive clusters with words with a syllable onset consisting of one consonant, in a case-study of one female child, aged 1:11- 2:5. The results showed a gradual progression through time in duration of the the syllable onset of target words with a voiceless plosive, but no clear trend in the duration of the target word with a complex onset. Czigler et al. (2000) noted that the complex onsets were all reduced to the voiceless, unaspirated plosive. They furthermore noted that the increase in duration of the target words with a single initial plosive was caused by the the child starting to produce the initial plosive with aspiration. Therefore, the difference between productions in this case,
might largely be explained by the emergence of contrast in aspiration. One difficulty of investigating /p/-~/s\P/ contrasts is that the full production sequence of the plosive is difficult to study acoustically in isolated productions due to the silent nature of the occlusion phase. The initial plosive may be less stable in duration in the child's production than the acoustic measurements show. The study presented here therefore investigated the contrast between the duration of syllable onset in monosyllabic targets words with a initial /sn/ or /sl/ onset and the corresponding word with a single element onset in a group of female speakers.

Method
Audio recordings of children were elicited using a set of target words in a mixed longitudinal and cross-sectional design form.

Speech samples
The target word set consisted of five Swedish words: 'sår' [soːr], 'lår' [loːr], 'slår' [sloːr], 'når' [noːr] and 'snår' [noːr].

Procedure
Eight female, monolingual Swedish children (77-149 weeks) were recorded at monthly intervals. The recordings were made to DAT tapes in a sound treated recording studio. In each session, the target words were elicited by the child's parent using hand drawn, black and white picture prompts. The children's parents were instructed to aim for an as spontaneous a production as possible, using primarily questions about the depicted object. As a secondary method, the parents were instructed to ask questions containing the word in non-final position (delayed repetition), such as “Is it a <target> or a car?”.

In the cases where the child failed to make a production using the delayed repetition method, the parents were instructed to try to elicit an direct imitation using questions containing the target word, e.g. “This is a [snoːr]. Can you say [snorː]?”.

Data inclusion criteria
A produced utterance was included into the analyzed dataset when a) the interaction between the child and the parent was consistent with the target word, b) the production did not co-occur with an external sound, and c) the production was made in a normal speaking style: prolonged, shortened, shouted and whispered productions were removed from the dataset.

Data analysis
The children's productions of individual words were extracted from the audio recording of the session. All recordings where the speech of the parent coincided with that of the child were removed. The remaining recordings were phonetically segmented and transcribed using WaveSurfer (Sjölander & Beskow 2000). The data were subsequently syllabified and extracted using the EMU speech database system (Cassidy & Harrington 1996). The statistical analysis were made using the R statistical package and the EMU-R interface.

Statistical analysis
Syllable onset was modeled using a mixed-effects linear model with 'onset structure' and 'Subject age' as independent variables. The factor 'Subject' was treated as a random effect. The model was fitted using the 'nlme' statistical package for R which does not depend on the dataset being balanced. (Pinhero and Bates 2002, p 25)

Results
Table 1 presents the frequency of occurrence in the collected dataset for each combination of subject and onset type. The total number of productions investigated was 667.

<table>
<thead>
<tr>
<th>Simple Onset</th>
<th>Complex Onset</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>69</td>
</tr>
<tr>
<td>F2</td>
<td>53</td>
</tr>
<tr>
<td>F3</td>
<td>60</td>
</tr>
<tr>
<td>F4</td>
<td>75</td>
</tr>
<tr>
<td>F5</td>
<td>133</td>
</tr>
<tr>
<td>F7</td>
<td>13</td>
</tr>
<tr>
<td>F8</td>
<td>67</td>
</tr>
<tr>
<td>F9</td>
<td>46</td>
</tr>
</tbody>
</table>
Inspection of the frequencies presented in table 1 shows that the groups created by the levels of the factor 'Onset structure' is not balanced.

A design plot of the dataset is presented in figure 1. The plot shows the mean syllable onset duration, grouped by the two onset complexity conditions, the investigated target word and subject. Figure 1 indicates that there is a tendency for the target words with a complex onset ([slo:r] and [sno:r]) to be produced with a higher mean onset duration than the targets with a simple onset ([no:r], [so:] and [lo:r]). Furthermore, the right bar of figure 1 indicates that there is a substantial variability in mean onset duration across subjects. Therefore, a single mean across subjects may not be assumed in the statistical modeling.

Table 2 presents an ANOVA table of a test of a mixed-effects model of produced syllable onset duration as described by subject age, with the category of the target word syllable onset structure as a fixed effect and subject as a random effect. The results from the ANOVA test showed a significant positive main effect of subject age and syllable onset structure category at the 0.01 level. No significant effect was found for the interaction of the two variables.

Table 2: ANOVA table showing the results of a test of a model of 'Onset duration' by subject age, with target word syllable onset structure category as a fixed effect and subject as a random effect.

<table>
<thead>
<tr>
<th></th>
<th>Num DF</th>
<th>Den DF</th>
<th>F-value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1</td>
<td>522</td>
<td>131.44</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Ostruct.</td>
<td>2</td>
<td>522</td>
<td>5.71</td>
<td>0.0035</td>
</tr>
<tr>
<td>Age:Ostr.</td>
<td>1</td>
<td>522</td>
<td>0.63</td>
<td>0.43</td>
</tr>
</tbody>
</table>

Figure 1. A design plot of the dataset. Mean syllable onsets durations are shown divided into groups by the factor target onset structure types (left bar), the investigated target words (middle bar) and subject (right bar).
Discussion and conclusion

The significant positive main effect of the variables age indicates that the groups of subjects are producing increasingly longer syllable onsets with age. Furthermore, target words with different structure in syllable onsets was shown to effect syllable onset durations differently.

However, the results of the statistical test presented in table 2 failed to show an interaction effect between the two variables. This is surprising since the results presented in previous reports from onset duration measurements have indicated a difference in the effect of onset structure on consonant duration at different ages. For instance, in the results presented in the case-study by Czigler et al. (2000), an effect of structure was only established at the later ages, not in the earliest stages of development.

One possible explanation for this difference may be that the effect may have been present also in the data of Czigler et al. (2000), even at early ages, but was manifested as articulatory variations of the silent interval of the occlusion.

In summary, the results presented in this paper suggest that there are acoustically measurable differences in duration between children's productions of a simple-target syllable-onset compared to those with a complex syllable-onset. Furthermore, the non-significant interaction effect suggests that the difference between the two syllable onset types does not vary significantly with age. The results therefore do not show any effect of the perceptual filter proposed by Spencer (1988).

The validity of the proposed explanation for the discrepancy between the data presented here and those in Czigler et al. (2000) should be investigated through comparisons of the development of the two types of consonants.

Acknowledgements

The author wishes to thank Kirk P. H. Sullivan for reading and commenting on earlier versions of this paper.

This study would not have been completed without the great effort put in by the investigated subjects and their parents in the recording phase of the study.

References


