

Margarita Gulian, Claartje Levelt
Leiden University
LIBC/LUCL

1. Introduction

Cluster reduction is a common phenomenon in young children production. In this paper cluster reduction in onsets will be discussed, where the second of the two segments is omitted, like in [dʌk] for truck and [si:p] for sleep. The goal of the study is to find out at which stage of the production process cluster reduction occurs. The idea is that a detailed analysis of the reduced form will help to determine the source of the deviation from the target form. For this purpose we compare children's productions of onset clusters that have been phonetically transcribed as reduced forms, to their productions of similar words that do not contain a cluster in the target adult form, by means of an acoustic analysis. The main finding in this study is that in the acoustic signal of reduced productions of target /r/-clusters there is a clear acoustic trace of the 'deleted' segment, providing evidence for a covert contrast between consonant clusters and singleton consonants.

To date there are several studies attesting covert contrasts in child language productions. In an overview article about covert contrasts studies Scobbie (1998) particularly mentions a study on onset clusters by McLeod et al. (1998, in Scobbie, 1998). Here, the VOT of child productions of [sk-] which was reduced to [k-] and the VOT of simple [k-] onsets was measured and appeared to be significantly shorter in the first case. This difference in VOT has been interpreted as a covert contrast that children make to distinguish between a reduced form and an original simple form.

In coda position too, significant differences have been found between reduced codas and the corresponding fully produced forms. In Song and Demuth (in press), a comparison was made between the reduced and the full production of words like 'dog' uttered by the same child (e.g.; [dɔ] vs. [dɔg]). They found that children used compensatory vowel lengthening in cases where the coda was reduced. Finally, also at a suprasegmental level covert contrasts have been pointed out, like in the study by Carter and Gerken (2004). Here, children had to repeat sentences like 'He kissed Lucinda' and 'He kissed Cindy'. In the cases

* This research was funded by NWO-Vidi grant 276-75-006.

where the children reduced ‘Lucinda’ to ‘Cinda’, there was a significant duration difference between the sentence containing reduced ‘Cinda’ and the sentence containing the full form ‘Cindy’.

Covert contrasts thus reveal knowledge that language learners have, but that is not made explicit enough for the listener in their productions. A form can thus be perceived as ‘reduced’, while the acoustic signal reveals presence of this apparently deleted material.

In the present study two commonly reduced clusters in Dutch child data, namely /r/-clusters and /n/-clusters, are acoustically analyzed and compared to corresponding words with singleton onsets. Thus, a target onset cluster /tr-/, produced as [t-], was compared to a phonetically similar word with a singleton onset [t-]. For instance, the utterance [tɛin] for *trein* (train) is compared to *tijd* [tɛit] (time). An example of the other cluster type is *knippen* /knɪpə/ (to cut), produced as [kɪpə] being compared to *kippen* [kɪpə] (chickens).

Before looking at the applied method in more detail, I will first provide some information on the theoretical background underlying this study.

2. Theoretical background

The main goal of the present study is to trace back the source of the cluster reductions in the child’s speech production mechanism. The speech production model by Levelt et al. (1999), capturing adult speech production is assumed for this purpose. According to this model (see Figure 1), speech production involves the step-wise retrieval of information and application of knowledge in different modules.

The production of a single word requires the activation of a lemma in the mental lexicon. Each lemma activates its corresponding word form, stored separately from the lemma. The word forms carry both metrical and sound specification information, mapped during the phonological encoding and syllabification phase. Subsequently, the phonological form is provided with an articulatory motor plan at the phonetic level, which is sent to the articulators.

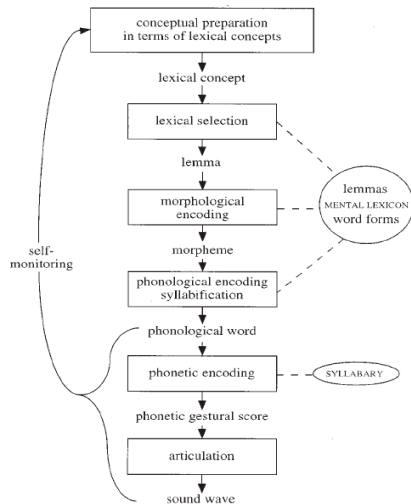


Figure 1: Speech production model of Levelt et al. (1999)

All the different stages between lemma selection and actual production are potential locations for cluster reduction. However, the reduction will have different characteristics depending on the locus of reduction in the model. If, for example, the target cluster has been stored in the child's lexicon with only one of the consonants, then we do not expect to find a trace of this unstored segment in the production. If we do find a trace, then, we can conclude that both consonants of the target cluster are present in the segmental representation, but the realization of the sequence of consonants is flawed, due to problems at lower levels of the production model. In this case, reduction occurs either at the level of phonological encoding, or at the level of phonetic encoding. At the level of phonological encoding, a constrained syllabification process could yield only simple CV syllables. However, this would probably also lead to a complete, i.e. trace-less deletion. A trace would thus most probably point to planning or timing problems with the execution of the motor plan for two consecutive consonants at the phonetic level.

3. Method

Data

Data from two different sources were analyzed. All data are from children acquiring Dutch as their first language. The main source consists of utterances from the CLPF database (Levelt 1994; Fikkert 1994). The CLPF corpus consists of longitudinal data of 12 Dutch-speaking children aged roughly between one and a half and two and a half years old. We used utterances of 6 of these children in the age range of 1;11-2;9. In addition, 7 toddlers were recorded at a Dutch day-care center in order to expand the data-set. The utterances of 3 of these children were analyzed,

since only these three reduced the two cluster types under study. The age of these 3 children ranged between 1;10 and 3;1.

Cluster types

Two cluster types were analyzed: plosive + /r/ onset clusters, like in the words *trein* (train) and *cracker* (cracker) and /kn-/ onsets, as in the words *knippen* (to cut) and *knoop* (button).

In order to identify the target utterance in the database a search was carried out in the Childes related Phon program (Rose, 2006) There a search aimed to find all utterances that contained a cluster in the target but missed the second consonant of that cluster in the phonetic transcription of the child's actual production. That is, target clusters like /tr/ and /kr/ followed by any Dutch vowel or diphthong were searched for, matching a child production that lacked the second cluster consonant, e.g. [te], [ta], [ke], [ka].

At the day-care center, toddlers were asked to repeat a list of Dutch words with initial clusters and real or non-words with a singleton onset that were matched to the cluster words (i.e. *knap* (clever) matched to *kat* (cat). The children's utterances were recorded with a Microtrack II digital recorder and a Rode NTG1 microphone. The utterances were phonetically transcribed by the author.

With respect to the /r/-clusters the present analysis is based on 27 word pairs, i.e. 27 target /r/-cluster words were matched to 27 target words with a singleton plosive onset. The selection criterion for the cluster words was that the second consonant (in this case /r/) was not represented in the phonetic transcription of the utterance. In other words, according to the transcription, reduced /r/-clusters were selected.

For the /n/-cluster words, 18 target cluster words were matched to 18 target words with a singleton /k/. The word pairs that were compared were always produced by the same child and, when possible, in the same session. Otherwise the pair came from recording sessions that were close in time.

Measures and results

/r/-clusters

As already mentioned, we looked at two types of onset clusters where the second consonant was omitted, namely /r/-clusters and /n/-clusters, comparing acoustically utterances like *trein* (train) to *tijd* (time) and *knippen* (to cut) to *kippen* (chickens). All the acoustic measurements

presented in this paper were made using Praat 5.0.10 (Boersma and Weenink, 1996).

Starting with the /r/-cluster type, a characteristic of /r/ is a very low F3 which results in F3, but also F2, of the subsequent vowel bending upwards (Kent and Read, 2002). Since it is sometimes difficult to discern the higher formants in the child's vowel productions, in the child productions only F2 was measured. More concretely, F2 was first measured at the immediate vowel onset (time 1), followed by a measurement at 1/4th of the entire duration of the vowel (time 2). In Figure 2 an example is given of a child production of the word *kraan* (faucet), where /r/ is produced and the raising of F2 can be clearly discerned.

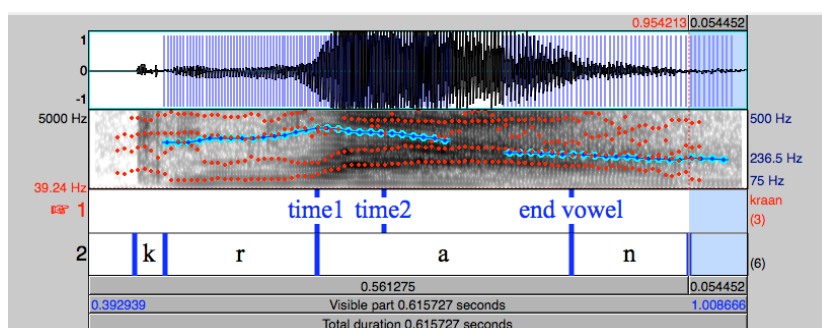


Figure 2: Waveform of the Dutch word *kraan* (faucet) produced by a two-year-old.

By defining the F2 values at time 1 and time 2, the potential raising of the formant, indicating a (trace of) a preceding /r/, can be captured. For this purpose time 1 is subtracted from time 2. A positive outcome indicates that F2 has bended upwards while a negative or zero outcome indicates a level (or lowered) F2:

- (1) time 2 - time 1 = positive outcome → presence of /r/
 2100Hz - 1800Hz = 300Hz → presence of /r/ measured in vowel /a/.

F2 was measured at these two points in the vowel both for reduced cluster utterances and their matching forms that contained no cluster in the target. Subsequently the attained outcomes of time 2 – time1 were compared by means of a paired sample t-test. In case the outcomes of the reduced clusters would not be significantly higher than those of the singleton onsets, the conclusion would be that the omitted /r/ does not leave a trace. In the opposite case, if the outcomes of reduced clusters would be significantly higher, the conclusion would be that a trace of /r/ is present in the data.

The vocalic formants of the 27 pairs were measured using a band filter analysis carried out in Praat. Band filter analysis of formants has been used lately by a number of scholars in the description of infant vowel productions (van der Stelt et al., 2005). The band filter analysis is a pitch-related analysis that is carried out using scripts in Praat. This analysis results in an estimation of a spectral envelope representation of an utterance. According to the inventors of this method of formant analysis (Wempe, 2001 in van der Stelt et al., 2005), it has an advantage above LPC¹ by being less sensitive to wrongly chosen parameters. Parameters are likely to be incorrect if there are high fundamentals in the data, which is often the case in child speech. For the formant analysis, the two points in time were selected manually, by the author, on the basis of the visual inspection of the waveform.

As was mentioned already, the outcomes of the formant values of reduced onset clusters were compared to these of singleton onsets. In this way the change in F2 in the vowel onset for the utterance [bo:t], where *brood* (bread) was intended, was compared to the change in F2 for [bo:t] where *boot* (boat) was intended. A paired sample t-test pointed out that the formant values for the reduced onset clusters differed significantly from the singleton onsets ($t(27)=2.97$, two-tailed $p=0.007$). Looking at the means of these two variables, 102 (difference in Hz) for the reduced clusters and -110 for singleton onsets, we can conclude that the F2 of the vowel is rising significantly in the case of the reduced clusters. In other words, although perceptively not present, acoustically a trace of /r/ was found in child utterances that were originally transcribed as having reduced clusters.

/n/-clusters

Let us now switch to the second cluster type, namely the /n/-clusters. There are two main properties of nasals that can be looked at when searching for a trace of nasality. The first property of nasals is nasal murmur, characterized by the presence of formants and an antiformant. The antiformant is a band of reduced energy. In addition, one of the formants characterizing the nasal murmur is the nasal formant, which is very low and comparable in energy to that of vowels (Kent and Read, 2002). Both the high-energy nasal formant and the energy-lacking antiformant can be measured through their bandwidth but they are also clearly visible on the spectrogram.

¹ LPC (linear predictive coding) is a standard and widely accepted method of measuring formants.

The second property of nasals is their F1, which is a separate formant, different from the nasal formant. Nasals tend to have an F1 at about 300 Hz which has an effect on the vowel, following the nasal. Therefore vowels like /i/, /u/ and /e/ tend to show a higher F1 than normal, while vowels like /a/ and /o/ will exhibit a lower F1 than normal.

The waveforms of the reduced /n/-clusters were first visually examined but signs of nasal murmur were not found. Therefore, next the F1 of the subsequent vowel was measured. Here too, reduced cluster forms were compared to suitable singleton onset utterances (*knippen* [kɪpə] to *kippen* [kɪpə]). Again, the band filter analysis was applied to measure F1. The mean F1 of the vowel in the two sets of utterances was measured and compared using a paired sample t-test. The statistical comparison did not show a significant difference between the two sets ($t(18)=1.21$, two-tailed $p=0.24$). This means that the F1 of the vowel following the reduced cluster did not turn to be significantly different (lower or higher, depending on the vowel) than the F1 of vowels following singleton onsets.

The 18 cases of reduced /n/-clusters as a group did not show any acoustic signs of nasality. However, looking at the individual waveforms of the utterances, those of the reduced clusters often do look different from the matching singleton onset utterances. In other words, investigating different acoustic aspects of a reduced cluster utterance like *knie* (knee) transcribed as [ki:] showed non-standard acoustic information that could be interpreted as compensation for the omitted nasal. Below, two concrete examples of these alternative productions of the onset cluster /kn/ are presented.

First, we will discuss the utterance transcribed [ki], produced by a girl at the age of 2;3, who is one of the children in the CLPF corpus. According to the database, while [ki] is produced, *knie* (knee) /kni/ is intended. The reduced cluster form is matched to the utterance [kɪkə] for *kikker* (frog) /kɪkər/. The singleton onset production is taken from the same session as the reduced cluster production. Looking closely at Figure 3, we can see that what at the first glance looks like a somewhat long /i/, actually consists of more elements than solely /i/. It is especially of interest that the higher formants become stable halfway what appears to be the vowel /i/. This suggests that /i/ is preceded by something else. Since perceptually, as a separate segment, this *something* sounds like a glide, this will be called the *glide strategy*, which is used to compensate for the omitted /n/ in the onset cluster. The mixed nature of the vocalic element

in [ki] becomes especially clear when compared to the vowel in *kikker*, which has a clear formant structure.

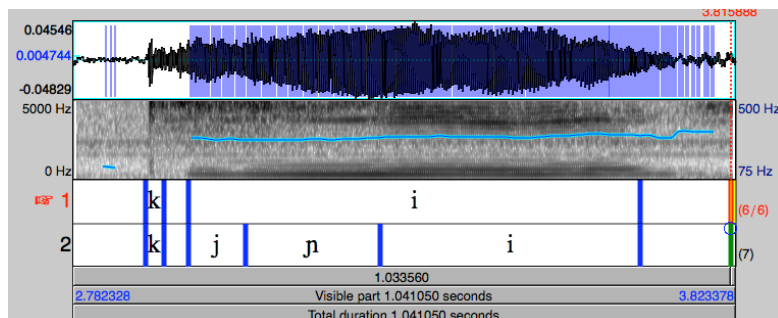


Figure 3: Waveform of the intended word *knie* which is transcribed as [ki] (Cato, 2;3 CLPF database).

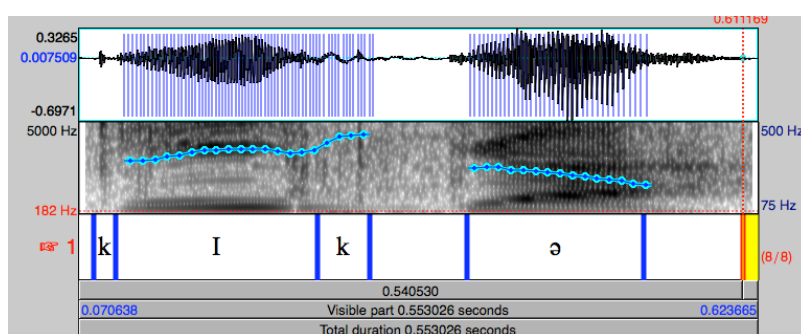


Figure 4: Waveform of the intended word *kikker* which is transcribed as [kɪkə] (Cato, 2;3 CLPF database).

The second example is the elicited production of the word *knap* (clever) /knap/ by a three-year-old. Perceptively *knap* lacks any form of nasality and is therefore transcribed as [kɑ:p]. *Knap* is matched to the production of the word *kat* (cat), transcribed as [kɑpf]. However, what strikes us again in the waveform is the different formant structure of the first part of what has originally been analyzed as the vowel /ɑ/ compared to the second part. A detailed analysis, guided by segmented perception, shows that the vocalic form starts with the vowel /a/ which then turns into the target /ɑ/. Here, it appears that a *vowel quality* strategy is used to compensate for the missing nasal. As can be seen, the vowel in *kat* has a more stable structure, even though its high formants are weakened.

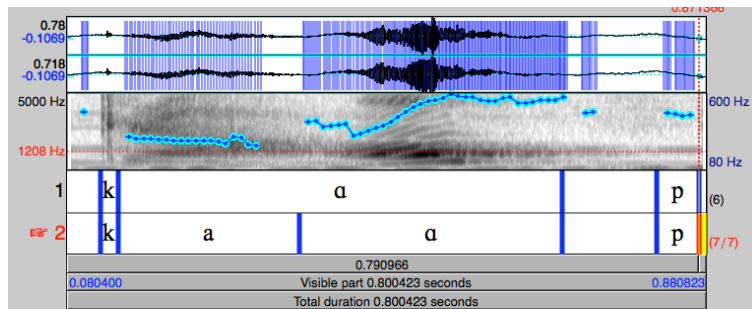


Figure 5: Waveform of the intended word *knap* which is transcribed as [kɑ:p], (Herman, 3;1 elicited production).

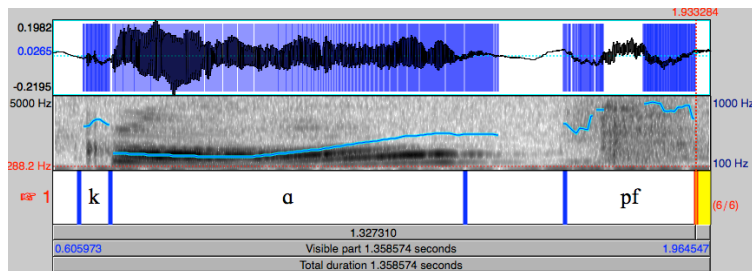


Figure 6: Waveform of the intended word *kat* which is transcribed as [kɑpf] (Herman, 3;1 elicited production).

There are more cases where the acoustic signal of reduced /kn/-vowel sequences deviates from the acoustic signal of singleton /k/-vowel sequences. However, the pattern is by far not as general as the compensatory cues we found for the reduced /r/-clusters. In future research, more data on the /kn-/ clusters will show whether it is possible find a pattern in the compensatory strategies for this type of cluster reduction.

4. Discussion

In the present study two different types of reduced clusters were analyzed, for which different results were obtained. In the case of plosive+/r/ onset clusters, a clear acoustic trace of the omitted /r/, namely a raising F2, was found in the vowel. However, in the reduced productions of target /kn-/ onset clusters, no acoustic trace of nasality – i.e. nasal murmur or a F1 lower or higher than usual, depending on the vowel – could be found. Although individual cases exhibited clear compensations for the omitted nasal, a general pattern could not be distinguished yet.

Although the findings are based on a limited number of utterances it is possible to draw some – tentative - conclusions about the possible source of these reduced clusters in the speech production mechanism of children. According to one of the predictions, if in an utterance with a reduced cluster the omitted segment leaves behind an acoustic trace, this entails

that in the child's mental lexicon the cluster is properly stored. Only then can the child intend to produce the segment. This, in turn, suggests that the actual reduction occurs at the lower levels of the production model. In the present study it was found that reduced /r/-clusters do indeed leave an acoustic trace in the productions of two to three-year-old Dutch children. It can thus be concluded that the onset cluster was stored properly. Now the question is whether this cluster is being reduced at the phonological or at the phonetic level.

If cluster reduction occurs at the phonological level – i.e. at the level of phonological encoding – syllabification in the Levelt et al. model – this could mean that there is a developmental constraint on the types of syllables that can be formed. If CCV(C) syllables cannot be formed at this developmental stage, then one of the onset consonants cannot be syllabified and will therefore be dropped. We thus expect a complete deletion of the second consonant at this level if syllabification is constrained to CV(C). This entails that no phonetic encoding of this segment can take place, so no trace of /r/ would be present in the signal. Since we do find an acoustic trace of /r/, it is unlikely that the source of the reduction lies at the level of phonological encoding. The acoustic trace shows in a way the children's intention to produce both consonants of the cluster. This, however fails.

The conclusion seems to be that the cluster is stored and retrieved from memory, it passes through the phonological encoding level but it is affected at the phonetic level. At this level the problem lies with executive control over the construction or the execution of a motor plan. Children at this stage probably have a limited capacity to plan and produce a complex sequence of segments within a certain amount of time. This leads to the partial realization of /r/.

In the case of reduced /kn-/ clusters no trace of an /n/ was found in the waveform. In individual cases some form of compensation was found, but this captured the quantity rather than the quality of the omitted /n/. There are two possible explanations that require further research.

One explanation is that a liquid is perceptually more salient for young children than a nasal in a consonant cluster. This could be due to the sonority distance between the two consonants, which is large for /r/-clusters but small for /n/-clusters (Selkirk, 1984). Storing the phonological details of /n/ in /n/-clusters could thus be harder than storing the phonological details of /r/ in /r/-clusters. One possibility is thus that the quantity but not the quality of /n/ is stored in memory, and that the phonetic implementation of quantity varies between children. The other

explanation is that /kn-/ is fully represented in the lexicon, but that at the phonetic level executive control over nasal quality, but not over the quantity of /n/, fails. More data and perception experiments will provide us with the answer.

5. Conclusion

In this paper two particular cluster types in Dutch child language were analyzed. More specifically, the reduced productions of these clusters were analyzed in detail by means of an acoustic analysis. The goal of this detailed analysis was to answer two questions: (1) can we find evidence for covert contrasts in the reduced clusters of young children, and (2) where in the speech production model does cluster reduction occur?

The acoustic analysis of the productions of reduced /r/-clusters showed that, even though according to their transcriptions they lacked /r/, these reduced clusters did leave a clear trace of the liquid in the waveform. This was different for the productions of reduced /n/-clusters, which did not show a trace of the nasal in the waveform. In individual cases of reduced /n/-clusters, it was found that children sometimes do compensate for the omitted nasal but in different ways, reflecting the quantity rather than the quality of /n/.

Two conclusions were drawn with respect to the source of the reduction in the speech production model. First, in the case of reduced /r/-clusters, the trace in the waveform showed that the cluster is stored in the mental lexicon and is phonologically encoded. The – partial - deletion of /r/ in the /r/-clusters most likely occurs at the phonetic level, due to problems with the executive control over the motor plan. Determining the source of reduction in /n/-clusters awaits further research. Two possible sources – the mental lexicon and phonetic encoding - have been identified, and well-designed perception experiments and more production data will lead to a definitive answer. Once again it has been shown that young children have more phonological knowledge than they can express in a way that is perceptible for adult listeners.

References

- Boersma, P. and Wenink, D. (1996), Praat: a system for doing phonetics by computer, version 5.0.10, Institute of Phonetic Sciences of the University of Amsterdam.
- Carter, A. and Gerken, L. (2004), Do children's omissions leave traces? *Journal of child language*, 31, 561-586.
- Fikkert, P. (1994), *On the acquisition of prosodic structure*. Doctoral dissertation. Holland Academic Graphics: The Hague.
- Kent, R. D. and Read, C. (2002), *The acoustic analysis of speech*. Thomson Learning: Canada.
- Levelt, C.C. (1994), *On the acquisition of place*. Doctoral dissertation. Holland Academic Graphics: The Hague.
- Levelt, W. J. M., Roelofs, A. and Meyer, A. S. (1999), A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1, 1-75.
- Rose, Y. (2006), Phon: a program containing corpuses of child speech data in the form of sound files, version 1.3r500, Memorial University, Canada.
- Scobbie, J. M. (1998), Interactions between the acquisition of phonetics and phonology, In: Gruber, C. et al. (Eds.), *CLS 34: The panels*, 343-358.
- Selkirk, E. (1984), On the major class features and syllable theory. *Language sound structure*. Eds. M. Aronoff and R. Oehrle. Cambridge, London: MIT Press, 107-136.
- Song, J. Y., Demuth, K. (in press), Compensatory vowel lengthening for omitted coda consonants: A phonetic investigation of children's early representations of prosodic words.
- van der Stelt, J. M., Zajdo, K., Wempe, T. G. and Pols, L. C. W. (2005). Exploring the acoustic vowel space in two-year-old children. Results from Dutch and Hungarian. *Speech Communication*, 47, 1-2, 143-159.