The development of the speech production mechanism in young children : evidence from the acquisition of onset clusters in Dutch

Gulian, M.E.

The handle http://hdl.handle.net/1887/57176 holds various files of this Leiden University dissertation

Author: Gulian, Margarita
Title: The development of the speech production mechanism in young children : evidence from the acquisition of onset clusters in Dutch
Date: 2017-10-31

Chapter 6. Discussion

The goal of the research presented in this thesis was to come up with a proposal for a developmental model of the speech production mechanism. Different aspects of the production and, to a smaller extent, the perception of target onset clusters by young children were studied in order to find out how the speech production mechanism functions in these developing speakers, and how it evolves over time. In this final chapter I will summarize the results from the different chapters in the light of the speech production model that was introduced in Chapter 1. I will use the results from the four different studies that were carried out in the thesis to present the state-of-the-art of what we have come to know about the developing speech production mechanism, and discuss issues that remain to be elaborated. Below I will start out by shortly outlining the model again that was introduced in Chapter 1. Then I will present the developmental view on the speech production mechanism that was proposed by Wijnen (1990) and Levelt (1998). Subsequently I will discuss what the results from the studies in this dissertation tell us about the development of the speech production mechanism, and to what extent they supplement the developmental perspective presented in the work of Wijnen and Levelt.

6.1. The model

6.1.1. Speech production

The speech production model that formed the point of departure for the studies in this thesis is depicted again in Figure 1, and incorporates the speech production model by Levelt et al. (1999) and the production chain of the bidirectional perception-production model presented in Boersma and Hamann (2009). The Auditory Target Form (Boersma, 2011) replaces the Phonetic Gestural Score in the model of Levelt et al. In the model as discussed here, speech (word) production involves the step-wise retrieval of information and application of knowledge from retrieval of the word form from the mental lexicon to articulation.

Lexical Entry

↓ lexical retrieval

Phonological Underlying Form

↓ phonological encoding

Phonological Surface Form

↓ phonetic encoding

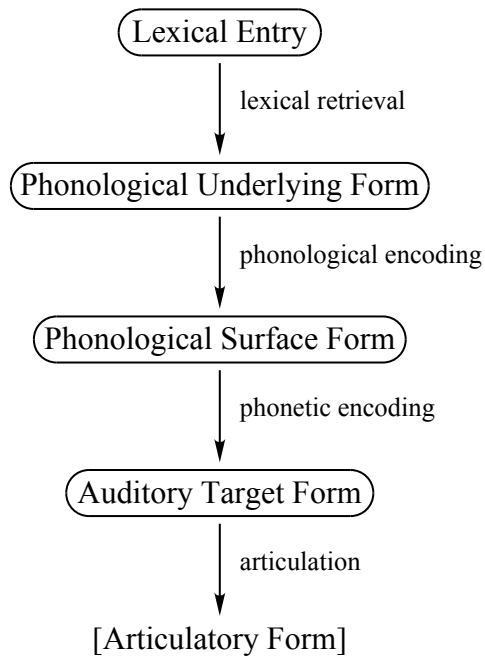Auditory Target Form

↓ articulation

[Articulatory Form]

Figure 1: The speech production model, based on Boersma and Hamann (2009), Boersma (2011) and Levelt et al. (1999).

Focusing on the word-form encoding part of speech production, the production of a word requires the activation of a lexical entry in the mental lexicon. If, for instance, a child intends to utter the word *schommel* 'swing', then the lexical entry <schommel> will be activated. Each lexical entry activates its corresponding underlying form, in this case /sxɔməl/, which contains the stored information about the word's sound segments. Phonological encoding of this information leads to a syllabified phonological surface form |sxɔ.məl|. Subsequently, phonetic encoding converts the surface form to an auditory-phonetic target form - with syllabic position-specific allophones [ˈsxɔməɫ]. This

is the auditory target form that the speaker aims to achieve - even if obstructions in the articulatory-motor system prevent the normal way of articulation, as bite-block experiments carried out with adults (MacNeilage, 1981, Gay et al., 1981) and with children (Oller & MacNeilage, 1983) have shown; even if speakers are articulatorily inhibited, they try to produce the word as close as possible to the form that is typically produced. This points to the existence of an auditory target form which a speaker aims to achieve in production. The auditory target form, finally, is transformed by sensorimotor knowledge to an articulatory-motor program that controls the speech muscles and results in the acoustic realization of the word [sʷxʷɔm·əɬ].

6.1.2. Speech perception

Although in this thesis the focus is on the development of the speech production mechanism, this cannot be done without taking into account the speech perception abilities of the developing speaker. In the initial step of word production, the sounds of the word to be produced are retrieved from the mental lexicon. These sounds have been stored in the mental lexicon through the speech perception system. The nature and actual content of the lexical representation is thus wholly dependent on the perception chain, which more or less mirrors the production chain. When a word uttered by a speaker is perceived by a listener, first of all the listener receives the acoustic signal. In case of the word *schommel*, , the acoustic signal [sʷxʷɔm·əɬ] is mapped onto the representation [ˈsxɔməɬ] at the auditory target level. The next step is mapping this form to its syllabified Surface Form (SF) /.sxɔ.məl./. Subsequently this form is mapped onto its underlying form |sxɔməl|. At this level, if necessary, information about the word's morphology is decoded. Finally, the phonologically underlying form (UF) is matched to its lexical entry <schommel>. In order to completely understand the forms of the word-productions that are uttered by the developing speaker, a parallel study of the developing perceiver would be necessary; if segmental material is lost or changed in the perception chain, this would lead to an incomplete or incorrect

stored form in the mental lexicon, which would entail that the production chain already starts out with incomplete or incorrect information. Currently it is hardly possible to perform longitudinal production and perception studies in tandem, because perception studies still rely on group results, while ideally comparison of individual patterns in production and perception would be needed to study their interaction in development. In this thesis only the study presented in chapter 5 gives us a peek into the relation between incomplete production and incomplete perception - or storage - of the sound structure of a word. I will come back to this study in section 6.3.1

6.2. The initial state of the production mechanism

Wijnen (1990) proposed that the production mechanism of the young language learner is a reduced, non-hierarchical and non-modular version of that of the mature speaker. Development entails 'layering': adding layers where specific processing takes place, thereby rendering the mechanism hierarchical and modular. Similarly, in Levelt (1998), the ontogenesis of spoken words was hypothesized to look like in Figure 2 below.
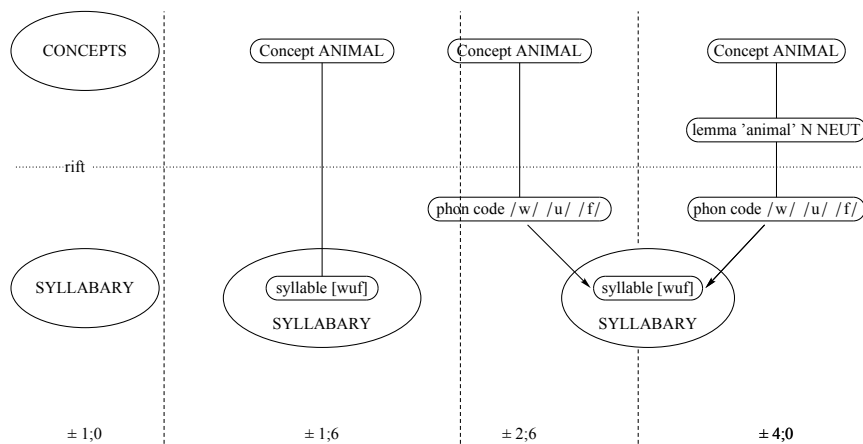
Figure 2: Four-staged model as suggested by Levelt (1998), visualizing how the child speech production model is built up.

The young speaker starts out with a reduced mechanism that consists of two as yet separate parts: a set of concepts and a set of articulations, i.e. syllabic babbles. At this point, concepts are not labeled in speech by any particular phonological form; there is no connection between concepts and articulations. The first meaningful words signal that a connection is established: concepts start to be labeled in production with relatively stable articulations, which are directly retrieved from the store of practiced syllables. At this point, then, words (or *protowords* according to Levelt) are not phonologically encoded. This layer becomes available when the ever-increasing number of words in the protolexicon forces the language learner to reorganize the system by *phonemization* - segmentalization - of the lexicon (Wijnen, 1990); between the ages of 1;6 and 2;6, words in the lexicon are fitted with increasingly detailed segmental representations. The establishment of the phonetic encoding layer is not separately discussed in Wijnen, nor in Levelt. However, since phonological encoding of the increasing number of segmental representations generates a series of newly formed syllables, phonetic encoding is required to generate

articulatory-motor programs for these new syllables, and establishment of this level is thus expected at this time. With experience these articulatory motor-patterns will probably be added to the syllabary. Around 4;0 the mechanism is expected to look like the adult mechanism (Levelt, 1998).

Since the focus of this study lies on the development of the production of onset clusters in words, the data that are studied result from a mechanism that has already developed the connection between concept and articulation, and works with at least partially segmentalized representations. This entails that both the phonological and phonetic encoding levels are already established. The data from the present study therefore mostly inform us about developments within these levels and the developing relationships between the levels, and they show us where the hurdles in production lie in the mechanism during development.

6.3. Sources of word production errors in young children

An articulated segmental error in the word production of a young language learner results from a deficiency in knowledge and/or processing ability during the word production process. The basic question in this study was whether characteristics of the error could point out the exact source of the error in the production mechanism. This, in turn, could then tell us more about the development of (specific layers in) the mechanism.

I will start out by discussing what our findings have revealed about the underlying form, i.e. the lexical representation of a word, the phonological encoding level and the phonetic encoding level, respectively.

6.3.1. Underlying Form

It is clear that the learner's articulated form will deviate from the adult articulated form if the underlying form, i.e. the sound sequence of a lexical entry stored in the mental lexicon of a young language learner, is incomplete compared to the adult target form, or contains an error. A possible source of the

reduced or substituted target onset clusters we encounter in young children's productions is thus the stored underlying form.

Deviating underlying forms would not only have an effect on production, but also on perception, since the perceived form has to be mapped onto the underlying form of a lexical entry at some point, resulting in a match or a mismatch with this form. In **Chapter 5** we therefore tested whether toddlers' underlying forms of target adult words with onset clusters matched better with reduced onset productions or with correct onset cluster productions, as revealed by their looking time to pictures of these target words upon hearing a picture label with a correct or a reduced onset cluster. Results differed for the two basic cluster types, /sC/ and /C+liq/. Correctly produced target /sC/ clusters lead to significantly longer looking times than reduced target /sC/ clusters, but this pattern ws not found for correct vs reduced target /C+liq/ productions. In both cases, the clusters were reduced to [C], which means that in target /sC/ clusters the initial consonant was omitted, while in target /C+liq/ clusters the second consonant was omitted - conform the initial productions of these clusters by young children. In this study, the majority of the toddlers who participated in the experiment reduced target /sC/ clusters in their own productions to [C], while they produced complete [C+liq] clusters for target /C+liq/ clusters. Since they produced [C+liq] clusters, these clusters must have been represented as such in the underlying form. The main question is, then, if their reduced target /sC/ clusters could have resulted from an incomplete lexical representation.

In similar experimental work where looking time to target pictures was measured in relation to segmentally correctly or incorrectly produced picture labels (among others White & Morgan, 2008) looking times to target pictures have been found to be shorter when labels were produced incorrectly. In these works, the conclusion is that the participants' representations are detailed and correct, and that a mismatching production therefore impedes lexical access,

resulting in the shorter looking times. However, with the knowledge from our study that the largest naming effect, i.e. increase in looking time, for *correct* productions was found for participants who reduced these clusters in their own productions, a different interpretation is possible: the *correct* production attracts attention, resulting in longer looking times, because it actually *mismatches* with the child's underlying form in a salient way, namely by the presence of [s]. This corroborates with the results of the Switch Paradigm study in Levelt (2012), where 14-month-olds who were habituated to the novel word [pa] showed significantly longer looking times when the switch test item [pat], containing an added [t] that could not have formed part of the representation constructed during habituation to the form [pa], was presented in the test phase. In both experiments, then, a mismatch between an underlying representation and a perceived form leads to longer looking times. Like in Levelt (2012), it can therefore be concluded that an incomplete underlying form can indeed initially be the source of an incomplete production.

Acoustical data from the longitudinal cluster production study reported in **Chapter 3** can be put forward as possibly supporting this conclusion. In that study, the productions of target /C+r/ clusters of five children were analyzed acoustically, from their first recording sessions up until the clusters were produced correctly, or if this did not happen, up until the end of the recording period. In **Chapter 2** it had been found that in a certain developmental period reduced cluster productions turned out to contain an acoustic trace of the omitted target /r/ in the following vowel. The presence of a trace of this /r/ in the produced form was taken to entail that /r/ had been present in the underlying form. In the longitudinal study in **Chapter 3**, however, it turned out that in an earlier stage, children did produce target C + /r/ clusters with a complete and traceless omission of target /r/. These traceless cluster reductions in production could very well be due to an incomplete underlying form, in which only a single consonant of the original onset cluster of the target word has been stored. However, in this case the form could also be due to

restricted phonological encoding, to which I will turn below. Perception data are necessary to distinguish between these two options. However, I could not check the performance of the traceless omitters in the study - and compare this to the performance of [+trace] omitters or correct producers - in a perception study with correct and reduced productions of target C+/r/ clusters because the production data had been collected between 1989 and 1991.

To conclude this section, the perception study from **Chapter 5**, in combination with the information on the production abilities of the participants in **Chapter 3**, shows that the underlying form is a likely initial source for the production of reduced clusters. However, in order to verify this, the combination of solid production and perception data is crucial. A study in which longitudinal production and perception data of a larger group of children is collected and compared is necessary to come up with a definitive answer on the role of the underlying form in the deviating forms of early word productions.

6.3.2.   Phonological encoding

During phonological encoding, the segments retrieved from the mental lexicon, i.e. the underlying form, are grouped into syllables. Syllabification takes place according to a combination of universal and language-specific rules. Universally, the most sonorant segment (usually the vowel) forms the nucleus of a syllable, and preceding consonants are preferably grouped into the onset, as long as the sonority sequencing principle is met. Remaining consonants are assigned to the coda position, or to an appendix position in the case of /s/ in /sC/ clusters (Trommelen, 1984). There are language-specific constraints on the complexity of syllables, specified in the phonological grammar. Some languages only allow CV syllables, i.e. they disallow codas and complex onsets, and require the presence of an onset consonant, other languages allow for complex codas but not for complex onsets, etc. (Blevins, 1995). It is well-known that children universally start out with a highly restricted syllable structure (Menn, 1976; Jongstra, 2003; Fikkert, 1994; Levelt, Schiller & Levelt, 2000). Levelt, Schiller and Levelt showed that Dutch children exhibit a systematic and

gradual development of syllable types in their production, starting with CV, depicted in Figure 3:

$$\nearrow \text{CCV(C)} \rightarrow \text{(C)VCC} \searrow$$
$$\text{CV} \rightarrow \text{CVC} \rightarrow \text{V} \qquad\qquad \text{CCVCC}$$
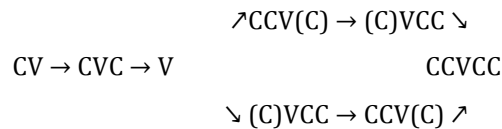$$\searrow \text{(C)VCC} \rightarrow \text{CCV(C)} \nearrow$$

Figure 3. Development of syllable types (Levelt et al. 2000).

If the syllable spell-out is constrained by phonological grammar, and the phonological grammar initially disallows complex onsets, the source for the omission of consonants from target adult consonant clusters could be situated at the level of phonological encoding. In the underlying form both consonants of a target onset cluster are represented, but due to the fact that only singleton onsets can be spelled out at the phonological encoding level, one of the consonants cannot be syllabified and ends up "stray"; it will therefore not be included in the Surface Representation. This entails that the stray consonant will be completely unavailable to the subsequent layer of phonetic encoding, predicting a traceless omission of the consonant in the child's production.

As discussed above, traceless omissions were encountered in **Chapter 3**, in the first stage of target onset cluster production. As long as clear additional perception data are unavailable, it is hard to determine which of the two possible sources, the underlying form, or phonological encoding, is responsible for traceless omissions in production. In the longitudinal data, however, traceless omissions of target adult /r/ soon occurred next to productions containing a trace, or even a more pronounced version of this /r/. The variable forms indicate that at this point , at least, the underlying form must contain the segment /r/.

With the longitudinal production study from **Chapter 4**, in which children were encouraged to participate in a Picture Naming (PN) task, a Word Repetition (WR) task and a Non-Word Repetition (NWR) task, I tried to argue from the correct and incorrect performance on the different tasks where the bottleneck in the production mechanism was situated at a given time, and how the bottleneck shifted over time. In the initial stage, a discrepancy was found between the performance on the NWR task and the performance on the PN task, where in the NWR task more advanced productions were encountered than in the PN task. Since picture naming requires both the retrieval of the underlying form and phonological encoding, while non-words in principle do not have a stored underlying form, and can skip the phonological encoding level, the more limited performance on the PN task must have been caused by the same possible sources identified for the spontaneous data from **Chapter 3**: either the underlying form, or phonological encoding. Again, we need perception data to be able to distinguish between the two options. Interestingly, however, both the spontaneous data from **Chapter 3** and the elicited production data from **Chapter 4**, from different children and collected in different decades, do show the same pattern: the main source for initial deviating word productions can be found in the upper layers of the speech production model: certainly at the level of phonological encoding. The exact role of the underlying representation remains to be determined.

6.3.3. Phonetic encoding

The Surface Form resulting from phonological encoding is mapped onto an Auditory Target Form at the phonetic encoding layer. Here, phonological features of segments are phonetically interpreted, and tailored to the segments' position in the syllable. Subsequently, an articulatory-motor program is constructed which will be executed by the articulators. In the adult model, there is evidence for a *syllabary*, where frequently used articulatory-motor programs for syllables are stored as entities (Levelt et al., 1999). Retrieving, rather than constructing articulatory-motor programs will speed up the

production process. In this thesis I did not specifically study the development of the syllabary, but this is certainly a relevant topic for future research, as it might show a U-shaped development; in the very first meaningful utterances, the beginning speaker seems to rely on a restricted set of ready-made motor patterns for syllables (see Figure 2 above), that are directly accessed. Subsequently these syllables are analyzed into their constituent segments, i.e. they are segmentalized (Wijnen, 1990; C. Levelt, 1994). The layers of phonological encoding and phonetic encoding develop at this point. With production experience, motor programs for syllables are stored as entities again. At this point, the syllabary thus seems to be reinstated - or reactivated - as part of the phonetic encoding layer. The syllabary could also play an important role in the variable forms of word productions we often find in child language, where in a single recording session more and less advanced realizations of the same target word can be found. Presumably, stored "old" motor programs for syllables compete with new, less frequently used new motor programs. I will come back to this below, in section 6.4, where I discuss variable forms.

Data in **Chapters 2**, **3** and **4** all point to the layer of phonetic encoding as the source of deviating - with respect to the adult model - productions. In **Chapter 2** it was found that upon acoustic analysis, productions in which target onset clusters appeared to have been reduced to singleton onsets, actually turned out to contain traces of the omitted consonant in the following vowel. We have to consider the possibility that the form with the trace is the form that the child has stored in the mental lexicon, which would entail that the source of the deviating production lies in the lexical representation, rather than at the phonetic encoding layer. However, in **Chapter 3**, a sequence of developmental stages was uncovered, showing a gradually increasing presence of the initially completely absent target /r/ in the production of target plosive +/r/ onset clusters. In my opinion, it is unlikely that the driving force behind these developments in production would lie in the constant updating of the lexical

representation of the target onset clusters. In addition, the production data, both the spontaneous data in **Chapter 3** and the elicited data from **Chapter 4**, showed quite some variable forms, which are also more likely to result from an unstable phonetic and/or phonological encoding layer than from variable lexical representations.

The presence of a trace of target C_2 in the child's production, or a vowel-like rendition of this target consonant, as found in the longitudinal data in **Chapter 3** does not only indicate unstable phonetic encoding however. When phonological encoding is still restricted to singleton onsets, the trace or vowel-like rendition of target C_2 can be seen as a means of expressing this target consonant within the limits set by the phonological encoding layer, i.e. a CV(C) syllable structure: in the absence of a C_2 position in the syllable, the target C_2 is spelled-out, for better or worse, in the nucleus position. In case the syllable is restricted to CV(C), both /r/ and the subsequent vowel share a single nucleus position, leaving very little room for the phonetic encoding of /r/; in this case we find just a trace of /r/ encoded in the vowel. In case phonological encoding can spell-out syllables with a complex nucleus, target /r/ can be phonetically encoded as a separate segment, albeit a vowel. In both cases, phonetic encoding is not the main source of the deviating production: it has done the best it could do with the phonologically encoded form.

It was determined that epenthesis, stage 4 of target /Cr/-cluster production development, is a genuine problem at the lower levels of the model, either phonetic encoding or the execution of the motor program. In this view, the vowel intervening between [C] and [r] in production is the acoustic result of an immature coordination of the two consonant articulations; the articulation of the second consonant is initiated after the release of the burst of the initial plosive consonant. Whether this articulation is due to the motor program, or to the execution of this program can at this point not be determined. However, subsequent consonantal substitutions of the target /r/ also point to problems

with the correct phonetic implementation of the features of this rhotic consonant at the phonetic encoding level.

In **Chapter 4**, finally, we saw a developmental shift in the performance of the different production tasks, from a better performance on NWR than on PN tasks, to a state where performance on all tasks was similar, but still not perfect. This, following Den Ouden (2002), indicates, that the main error locus shifts from a location either at the level of the lexical representation or the phonological encoding level, to a location at the level of phonetic encoding. Here too, then, the data in **Chapters 3** and **4** corroborate each other in showing that the main source of deviating productions at later stages can be found at the lower, phonetic encoding layers of the model.

6.4 Variable forms

What do variable forms say about the developing speech production mechanism? Of course mature speakers produce variable forms too, depending for instance on speech rate, whereby higher speech rates can lead to segmental reductions and assimilations. In this thesis, the link between speech rate and variability between more complex and more reduced forms was not specifically studied. However, the production study in **Chapter 4** explicitly relied on variability in production success in different types of production tasks - picture naming, word repetition and non-word repetition - to infer the source of difficulties in the speech production mechanism. For example, success in cluster production in the word repetition task, combined with failure in the picture naming task indicates that the source of difficulty lies at the level of phonological encoding, since this level can be by-passed in word repetition, but not in picture naming. Variable forms thus form an important source of information on the developmental state of the mechanism.

In the study of longitudinal spontaneous data in **Chapter 3**, it turned out that despite the fact that a specific sequence of stages could be discerned in the

development of cluster production, these stages also showed considerable overlap. In these data, variability was not related to different production tasks, and the only possible context for more versus less correct productions was found in the data of Cato, where the same target word was produced correctly - with a cluster - utterance-initially, but incorrectly in non-initial positions in the utterance. Only a very small set of data was available, but it is worthwhile to study this possible context for variability in an experimental set-up in future research, as it might point to a role for attentional resources. For now, the overlapping stages show us what the upper limits on the form-encoding abilities of the production mechanism are. In addition, they reveal the relative instability of the newly developed abilities; for reasons that we will need to study experimentally, the new abilities cannot always be used to the full, and the developing speaker then reverts to 'old' well-practiced spell-outs or articulatory-motor programs.

6.5 The developing speech production mechanism

At the time when children produce reduced onset clusters, which is the focus of interest of this dissertation, the speech production mechanism resembles the adult model in its complexity. This means that the imperfections visible in the child productions that were studied in this thesis are not due to an incomplete mechanism, with missing layers, as proposed by Wijnen (1990) and as depicted in the first four ontogenetic stages proposed by Levelt (1998), above in Figure 2. When a target onset cluster is realized in a way deviating from the adult model realization, this must then either result from an immature representation at one of the levels of the model, or from a problematic mapping from one layer of the model to the next. The variable forms we encountered in the longitudinal data in **Chapters 3** and **4** point to the fact that the speech production mechanism is still unstable, but they also help to determine the error source more precisely; in the case of variation it is more likely that the representation at some level is correct, while the problem lies in the mapping between that level of representation and the next-lower level.

According to the current account, then, the speech production mechanism matures in a top-down fashion. Constraints on syllable structure initially limit the segmental spell-out options at the level of phonological encoding. Phonetic encoding thus operates on a phonologically encoded form which already deviates from a form encoded by a mature production mechanism. Phonetic encoding can result in additional deviating characteristics in the form - compared to a form phonetically encoded by a mature mechanism - that will be articulated.

The produced word form, at every developmental stage, thus results from phonological encoding, which operates within the limits set by the developing phonological grammar, and phonetic encoding, which translates this phonologically encoded form into an articulatory plan - and is constrained by its own developmental limitations. This interaction can be illustrated with the description of developmental stages in the production of target /Cr/ clusters.

Stage: /r/ is completely omitted:
Phonological encoding: syllable spell-out is constrained by the phonological grammar, which initially does not allow for complex onsets (see Fikkert, 1994; Demuth & Fee, 1995; Levelt, Schiller & Levelt, 2000).
Result: /r/ is not encoded: becomes stray segment
Phonetic encoding: Stray segments are not seen by this module.
Result: target /r/ is not included in the articulatory-motor program

σ
Rhyme
Onset   Nucleus   Coda
b r   u   k
"stray"
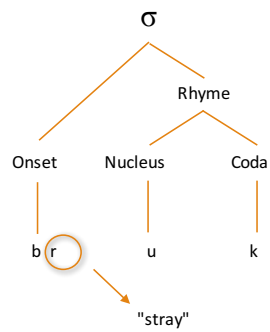
Stage: /r/ = F2 trace

<u>Phonological encoding</u>: syllable spell-out is still constrained by a phonological grammar that does not allow for complex onsets. However, there is a new development: (grammatical) pressure to spell-out all underlying segments, leads to /r/ being spelled-out in the Nucleus.

<u>Phonetic encoding</u>: the now phonologically encoded |r| shares single Nucleus position with |u|, and can only be accommodated very partially.

Result: |r| is phonetically spelled-out as F2 raising in the vowel.

σ
Rhyme
Onset   Nucleus   Coda
b   r u   k
F2 trace in [u]
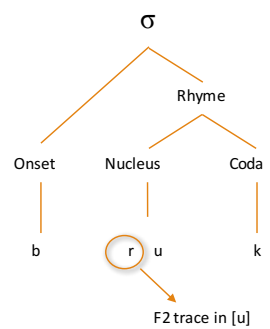
Stage: /r/ = full vowel

Phonological encoding: the phonological grammar now allows for a complex nucleus (but not for complex onsets). Both /r/ + /u/ are spelled out in a complex Nucleus.

Phonetic encoding: |r| is now phonologically encoded in an independent Nucleus position, and can be more fully accommodated phonetically.

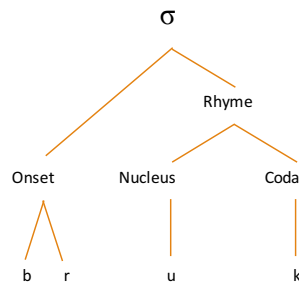Result: |r| is phonetically encoded as full vowel
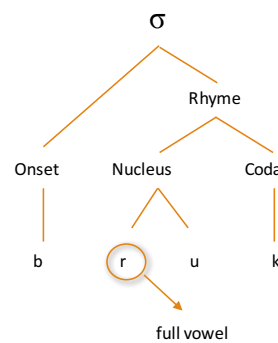
Stage: epenthesis in target cluster

Phonological encoding: there is a new development in the phonological grammar, which now allows for complex onsets.

Result: Both /C/ and /r/ are spelled out in a complex onset

Phonetic encoding/articulation: the coordination of the two consonants is either not correctly planned or not correctly executed: [b_ruk]

Result: Articulation of the second consonant is initiated after the release of the burst of the initial plosive consonant (Gafos, 2002), which results in a perceived epenthetic vowel

```
                    σ
                   / \
                  /   \
                 /   Rhyme
                /    /  \
               /    /    \
          Onset  Nucleus  Coda
            /\      |      |
           /  \     |      |
          b    r    u      k
```

As soon as the articulatory coordination of the two consonants in the complex onset position is fixed at the lower levels of the speech production mechanism, the production of a target /Cr/ cluster will be correct.

6.6 Conclusion

In this thesis I have tried to gain insight into the developing speech production mechanism by studying young children's production and perception of target words containing an onset cluster in different ways, and by using different methods. For future research it seems especially worthwhile to make use of the method based on Den Ouden (2002), which was used to discover the sources of production errors by combining the results of different production tasks. Ways should be found to add the perception task, which was missing here. In general, perception tasks are a necessary addition to production tasks in order to be able to differentiate between problems at the level of lexical representation and problems at the level of phonological encoding. Another valuable source of insight in the present work was formed by the acoustic data, as they were able to show both the presence of segmental knowledge that would otherwise have remained unacknowledged, and the ways in which the speech production mechanism spelled-out these stored segments. The developmental model I presented in the end is limited by the data I was able to collect and analyze - which always forms a major challenge for researchers in the field of child language. However, I hope to have shown that using a combination of methods

and data is necessary to understand how the speech production mechanism develops, and that the model constructed on the basis of these data is a model that can be further explored and tested.