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

Gulian, M.E.

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

The development of the speech production mechanism in young children:

Evidence from the acquisition of onset clusters in Dutch

Margarita Gulian

Published by
LOT phone: +31 30 253 6111
Trans 10
3512 JK Utrecht e-mail: lot@uu.nl
The Netherlands http://www.lotschool.nl

Cover illustration: Meike Fortuin performing a pilot perception test at home.

ISBN: 978-94-6093-257-1
NUR 616

The development of the speech production mechanism in young children:

Evidence from the acquisition of onset clusters in Dutch

Proefschrift

ter verkrijging van

de graad van Doctor aan de Universiteit Leiden,

op gezag van Rector Magnificus prof. mr. C.J.J.M. Stolker,

volgens besluit van het College voor Promoties

te verdedigen op dinsdag 31 oktober 2017

klokke 10.00 uur

door

Margarita Etvart Gulian

geboren te Sofia, Bulgarije

in 1981

Promotores: prof. dr. C.C. Levelt, Universiteit Leiden

prof. dr. N.O. Schiller, Universiteit Leiden

Promotiecommissie: prof. dr. J.C. Schaeffer, Universiteit van Amsterdam

prof. dr. F.N.K. Wijnen, Universiteit Utrecht

prof. dr. L.C.J. Barbiers, Universiteit Leiden

dr. B.M. van 't Veer, Universiteit Leiden

This thesis is part of prof. dr. C.C. Levelt's VIDI project "A psycholinguistic model for language acquisition" project number 276-75-006, financed by NWO.

Table of contents

Acknowledgements

I would like to thank my supervisors Niels and Claartje for helping me finish up this project called PhD thesis. I would especially like to thank Claartje for helping me rewrite a big piece of the thesis and being so patient with me. I would also like to thank the committee for their useful comments.

I want to thank the children recorded for this thesis and their parents, especially the children recorded for chapter 4, Hannah, Matteo, Maaike and Lars and their parents.

In one breath, I would like to thank Paul and Caroline for co-authoring chapter 2 and chapter 5, respectively. You helped me give a better presentation of my data. Besides, I want to thank Mirjam de Jonge and Monique Bisschop for being the best lab assistants ever. Last but not least, I want to thank Iris for helping me with the layout of the thesis, thank you for your patience!

I want to give a special thanks to my paranymphs Katja and Jan-Willem, in their role of guarding angels.

Here I would like to thank a bunch of other people, with who I maintain a relationship in one way or another. Since I would hate to rank my relationships to these people I have decided to randomize their names. I am grateful to you all!

I want to thank the ones who helped with the statistics in the thesis, the ones who helped thinking how to order chapter 4, the one with who I worked together on our theses in Zeeland, the ones who joined me to concerts, the ones who danced with me, the one who accompanied me on guitar, the ones who sung with me, the ones who taught me new things, the ones who went jogging with me, the ones who made me feel welcome at the UvA, the one who taught the Leiden PhDs how to play hacky, the ones who patiently shared an office with me, the ones who wrote me long e-mails from the other side of Europe or from another continent, the ones who inspired me with new music, the ones who took care of my kids, the ones who think of me, the ones who love me: Vidhi, Aude, Janitsa, Margarita, Marthy, Irina, Jessie, Marieke, Bilyana, Iviana, Elena, Wolfgang, Tsvetan, Erik, Victoria, Gideon, Roberta, Annemiek, Catherine, Petrus, Paula, Dimitar, Annelies, Roland, Sita, Luz, Irene, Rosa, Aura, Sevda, Meba, Anne, Jurriaan, Charlie, Jos, Robin, Marijn, Linda, Sara, Wieneke, Karin, Elitsa, Arnoud, Maria, Yimmy, Serge, Elitza, Frank, Eti, Marcela, Roman, Nana, Fang, Zheni, Pieter, Angélica, Allison, Ineke, Kathrin, Dessi, Robert, Teo, Rebecca.

A special thanks goes to my parents in low for all their help, literally with everything, superveel bedankt Hennie en Anneke!

I also want to thank my parents, Rossi and Edi, my brother and sister, Tani and Paola and my living grandma, baba Tinka, but of course from somewhere far away baba Pepi is constantly supporting me. Нямам думи да ви се отблагодаря за обичта ви!

En natuurlijk mijn gezinnetje, Martijn, Kalina en Radana. Страшно ви обичам и не знаете колко е хубаво, че сте част от живота ми!

Chapter 1. Introduction

1.1. Introduction

This thesis is about children's developing word production skills, and about the development of the system behind language production. The production of speech by adults has been studied in great detail, leading to several different models of the processes involved (Dell, 1986; Levelt, 1989; Levelt, Roelofs & Meyer, 1999; Boersma, 2011). However, up until now this line of research has hardly ever been extended to the (typically) developing speaker (cf. Wijnen, 1990; Stackhouse & Wells, 1997; Levelt, 1998). Despite the fact that child language productions typically deviate from the adult standard, the way the speech production mechanism performs and develops in the early stages of language production is largely unknown. In most work on phonological acquisition to date, some developmental state of the child’s grammar is held responsible for these specific productions. However, the child language data that are studied are always production data; ignoring the real-time processes that have shaped these productions yields an incomplete account of the data (Docherty & Foulkes, 2000). We thus need to know more about the speech production mechanism of the developing speaker, and with this thesis I hope to contribute to this call.

I have limited the work in this thesis to a study of the system behind the production of isolated words, since this is what the developing speakers in this thesis, being between one and two-years old, mostly produce. Within the context of word-production, this study will focus on the - developing - production of word-onset consonant clusters. A typical deviation in early child language productions is the reduction of these clusters to singleton consonants, like in (Dutch) [tɛin] for target *trein* ‘train’, and [tup] for target *stoep* ‘side-walk’. As mentioned above, up until now we only find grammatical accounts of this deviation, in the form of a fixed syllable template, a parameter setting, or a constraint on syllable structure (Fikkert, 1994; Pater & Barlow, 2003; Velleman

& Vihman, 2002). A brief discussion of these accounts will follow below in 1.4. However, instead of resulting from a specific grammatical setting, these cluster reductions could also be the outcome of the speech production process, and in the speech production mechanism there are several possible sources for error that could be considered. This is what will be done in this thesis, by studying children's cluster productions in different ways - acoustically, phonologically, and in relation to children's perception of consonant clusters - and analyzing both longitudinal, spontaneous production data, and elicited productions.

1.2 The speech production mechanism

The different possible sources for error in child language productions that will be studied are the layers in the model depicted in Figure 1, based on the speech production model of Levelt, Roelofs and Meyer (1999) and the bidirectional model of Boersma and Hamann (2009), and Boersma (2011).

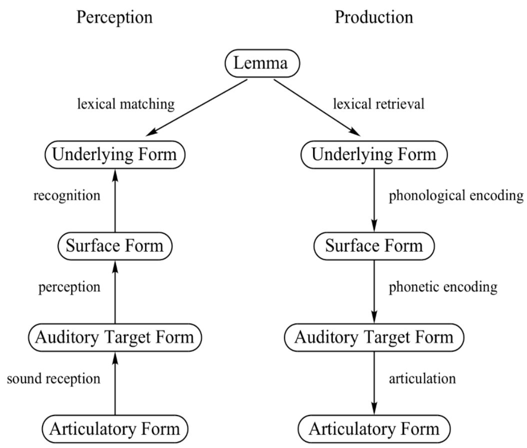


Figure 1. The perception-production model used in this study, elaborated on the basis of Boersma and Hamann (2009) and Levelt et al. (1999).

According to this serial processing model, and focusing first on the production side, in the mind of a speaker an intended concept is transformed in several steps into a motor program that will eventually be executed by the articulators. It takes around 600-700 ms from the moment of seeing a picture of a common object, like a train, to the moment of uttering the monosyllabic word *train* in a picture-naming task (Indefrey & Levelt, 2004; Szekely et al., 2004). In this very short time, the following steps have taken place:

1. Lemma activation (lemma = non-phonological part of an item's lexical information; Levelt, 1989). In the case of *train*, the lemma <train> will be activated.

2. Lexical retrieval. Each lemma activates its corresponding underlying, morphologically encoded, phonological form, which contains the stored information about the word's sounds, in this case /tren/, and the metrical frame, i.e. the number of syllables and stress pattern.

3. Phonological encoding. From this information, a phonological surface form is created. At this level, sounds are grouped into syllables, a single one in the case of *train*. I assume that this happens in a top-down way: segments are mapped onto stored syllable templates.
4. Phonetic encoding. Subsequently, the surface phonological form is converted into an auditory target form. In Levelt et al. (1999), it is assumed that for experienced speakers, motor programs for fequently-used syllables are stored in a mental syllabary, and can be retrieved directly. If a ready-made program (or the syllabary as a whole) is not available, the surface phonological form is provided with position-specific articulatory detail on the fly. In Levelt et al., the result of phonetic encoding is called the phonetic gestural score, but in Boersma and Hamann (2009) the phonetic encoding part is worked out in more detail, and is split into two modules, one that maps the surface phonological form onto an Auditory Target form, and one where this form is mapped onto an articulatory-motor program. Bite-block experiments have shown that speakers intend to produce vowels as closely as possible to an acoustic target, even when production is articulatorily inhibited (MacNeilage, 1981). This points to the existence of an auditory target form, which a speaker aims to achieve in production. The auditory target form is subsequently translated into an articulatory-motor program that controls the speech muscles. However, due to the limits of the present study, in this thesis, like in Levelt et al., a single phonetic encoding module is considered as possible error locus. Here, the phonological surface form is converted into the motor action instructions that will result in a form that the speaker aims to achieve in production, i.e. the auditory target form.
5. Articulation. The auditory target form is executed by the articulators, resulting in the acoustic realization of the word: [tɹẽːn]

Although the main concern of this thesis is the speech production system, we need to take perception into account too. Speaking can hardly do without perceiving, decoding and representing speech. The model in Figure 1 includes this component. For word production, the focus of this study, the speech comprehension system does not only play a crucial role in the way the sounds of words are stored - if certain sounds are not stored, they will certainly not be produced either - but also in what is called 'self-monitoring' by the speaker during the production process. Speech is monitored by the speaker before it is overtly articulated, as soon as a phonologically encoded form is available. For self-monitoring, the perception part of the model is used by the speaker, i.e. self-perception of inner speech takes place. If necessary, namely when an error is detected, repairs can be made before the speech is uttered. In the present study, I focus on perception only in relation to the segmental representations that form the input to the form-encoding part of word production. However, for a full understanding of the way developing speakers produce speech, perception and production and the systems underlying these processes should be studied in tandem. My hope is that as a sequel to the present work, the full model as depicted in Figure 1 above, will be studied in relation to phonological development.

1.3. Different sources of cluster reduction

For the developing speaker, like for the mature speaker, all the different stages between lemma selection and actual articulation are potential locations for error, resulting in productions that deviate from the standard. For this study, it is assumed that the exact source of the error in the production mechanism can be deduced from the type of error that results. This, in turn, can inform us about the developmental state of (specific layers in) the mechanism.

If, for example, the target cluster is incompletely stored in the child's mental lexicon, with only one of the consonants, the error source is the underlying form, i.e. the segmental representation. In this case, we expect to find a highly systematic error; the consonant that is absent from the representation cannot be encoded in any way, so there will be a systematic and complete omission of this segment in the speaker's production. If, however, a target cluster is variably produced correctly and incorrectly, we can conclude that both consonants of the target cluster are present in the segmental representation. An incorrect realization is then due to problems at lower levels of the production model, either at the level of phonological encoding or at the level of phonetic encoding. A single type of data is in general not enough to determine the exact error locus, and a combination of informative data needs to be considered. In Den Ouden (2002), an inspiration for the present study, the error locus in the production mechanism of patients with aphasia was determined on the basis of their performance on three different tasks, Picture Naming, Repetition, and Phoneme Detection. Arguing from the combined results of success on one task and failure on another, Den Ouden deduced whether the weakest link in the mechanism was formed by lexical access, phonological encoding or phonetic encoding. In Chapter 4 of this thesis, a similar procedure is used to find out about the development of the production mechanism.

1.4. Phonological accounts of cluster reduction

In phonological accounts of cluster development, usually two basic developmental stages are posited: an initial stage in which the underlying form /C_1C_2/ is reduced to a singleton [C] in the surface form – most commonly to C_1 if the target cluster consists of an obstruent followed by a sonorant - and a second stage in which a complete cluster can be present in the surface form, either correctly or with substituted segments. The initial stage, in which the cluster is reduced to a single C has been accounted for in different ways, and I will discuss the three most common ways here.

Template account. In this type of account, the child's production is constrained by a fixed template onto which consonants and vowels are mapped. Initially, this template is the core syllable, CV (Menn, 1976; Demuth & Fee, 1995; Demuth, 1996). An underlying representation /tren/ that is mapped onto this CV template, will end up as [te] in the surface form - and subsequently in production - because there are no positions available for the segments /r/ and /n/ in the template. This is shown in Figure 2.

Underlying representation: /t r e n/

Template: C V

Output: [te]

Figure 2: Cluster reduction in a Template account

Parameter account: Following the work by Chomsky (1981), Dresher and Kaye (1990) proposed a set of parameters governing the metrical structure of language. With respect to syllable structure, languages differ in their settings of parameters like the Minimal Onset Parameter ("Are Onsets obligatory?") and the Maximal Onset Parameter ("Can onsets be branching?"). In the initial stage of development, all parameters are in their default setting, and by paying attention to the input, the language learner will be able to change the default setting to the marked setting if evidence for this setting is present in the input. The default value for the Minimal Onset Parameter is *yes*, while for the Maximal Onset Parameter it is *no*. Together, these settings result in an initial grammar which only allows for syllables that have a single, obligatory consonant (Fikkert, 1994). In this initial stage, then, consonant clusters cannot be realized.

Optimality Theory account: In Optimality Theory (Prince & Smolensky, 1991), the phonological surface form results from an interaction of Markedness constraints, enforcing well-formedness, and Faithfulness constraints, enforcing the unaltered presence of information provided by the underlying form. The ranking of these constraints in a grammar determines the ultimate surface form of a specific underlying form. In the initial stage of development, Markedness constraints outrank Faithfulness constraints, and surface forms will thus have an unmarked, or well-formed, structure. Markedness constraints on syllable structure are Onset ("A syllable should have an onset"), No-Coda ("A syllable should not have a coda"), No-Complex-Onset ("A syllable should not have a complex onset") and No-Complex-Coda ("A syllable should not have a complex coda"). Only CV syllables can be the output of the initial grammar, where all markedness constraints are ranked high (Gnanadesikan et al., 1995; Levelt, Schiller & Levelt, 2000).

In all three accounts, the phonological grammar enforces complete omission of one of the cluster consonants in the initial stage, and complete onset consonant cluster realization in the surface form in a subsequent stage, if required by the underlying form. Depending on the theory, development leading to the subsequent stage consists of the availability of a new template, CCV, the Maximal Onset Parameter setting changing from default *no* to *yes*, or a demotion in the ranking of the constraint No-Complex-Onset with respect to a Faithfulness constraint, allowing for violations of the markedness constraint. There are, thus, no intermediate forms of a target cluster in a grammatical account. In Chapters 2 and 3, however, we will encounter data that are difficult to explain in a grammatical account because the C_2 is neither completely absent, nor completely present, or variably present or absent.

If we try to reconcile the phonological accounts with the psycholinguistic model, and a with a word production account, we could say that a grammar actually describes the limitations on the syllabification process in the phonological encoding module. This entails that if the problem with cluster realization lies in the phonological encoding module, we can expect complete, i.e. trace-less omissions of the underlying cluster segment C_2 because there is no position for this consonant available in the syllable inventory that can be employed by phonological encoding. When we encounter data like in Chapter 2 and 3, where the target C_2 is neither completely absent from, nor completely present in production, these are thought to result from flaws in the phonetic encoding module, or from a specific interaction between phonological and phonetic encoding.

1.5. Data

In this thesis, I have used both spontaneous and elicited data. In addition to studying production data, I carried out one perception experiment with young children (Chapter 5) and one with adults (Chapter 2).

The spontaneous word productions that I studied for Chapter 2 and Chapter 3 come from the CLPF database (Fikkert, 1994; Levelt, 1994) and are available through the CHILDES/Phonbank online database http://phonbank.talkbank.org/ (Rose et al., 2006; Rose & MacWhinney, 2014). The CLPF corpus consists of spontaneous speech production data, of 12 children between 1 and 2 years of age at the start of a one-year data-collecting period, acquiring Dutch as their native language.

In addition, for the study in Chapter 2, I recorded 30 children with a mean age of 2;1 years at four Dutch day-care centers in the Amsterdam area, and for the study in Chapter 4, I used longitudinal data collected from four children who were between 1;7 and 2;1 years old at the start of the data collecting period in the Amsterdam area.

For the perception experiment in Chapter 2, thirty-five adult speakers of Dutch were tested, in order to find out whether they were able to discriminate reduced onset clusters from singleton onsets, produced by Dutch two-year-olds. For the perception experiment described in Chapter 5, fifty-eight children with a mean age of 2;0 were tested.

More specific information about the participants in every study is provided in the separate chapters.

1.6. Overview of the thesis

The study in **Chapter 2** concerns the question whether reduced clusters in children's productions are indeed fully reduced - warranting a phonological account of cluster reduction - or whether they exhibit acoustic traces of the omitted second consonant. For this purpose, the acoustic characteristics of pairs of utterances, produced by the same speaker and at the same age, are compared, that differ only - or mostly - in the presence or absence of an onset cluster in their target forms, like *brood* /bʀoːt/ 'bread' – *boot* /boːt/ 'boat' and *knip* /knɪp/ 'cut' – *kip* /kɪp/ 'chicken'. These words are realized in such a similar way that even trained phoneticians tend to transcribe them identically, e.g. as [boːt] – [boːt] and [kɪp] – [kɪp]. An acoustic analysis of these forms, however, reveals acoustic traces of the omitted consonants from the target clusters in the children's productions. The children in this study tended to realize a rising F2 in the vowel onset when the target C_2 was /r/, which might be reminiscent of the rising F3 that we see in adult speech. As for target words starting with /kn/, where /n/ was omitted from the production, we found that the subsequent vowel did show a moving formant pattern, and a lower center of gravity. In a subsequent perception experiment with adults, where they were presented with these semi-reduced utterances and their minimal pair counterparts, it turned out that these adult listeners could not decide which of the two productions referred to a target word starting with a consonant cluster.

In **Chapter 3**, we take a detailed look at the acquisition of clusters starting with a plosive and followed by /r/- hence /Cr/ - over time, by five different children, in their spontaneous speech. All their attempts to produce target /Cr/ clusters, from the start of the recording period until the cluster is produced correctly - or until the end of the recording period - are analyzed acoustically. Although the five children show individual developmental paths, a general pattern can be discerned; in Chapter 2 partially reduced clusters were found, here it is found that this type of realization forms a developmental stage, preceded by a stage in which complete omission of the C_2 takes place, and followed by stages in which the C_2 becomes more and more present and then becomes more and more correctly realized. The different stages are discussed in terms of developments in the speech production mechanism.

In **Chapter 4**, we look at the longitudinal performance of four children on three production tasks: PN (picture naming); WR (word repetition) and NWR (nonword repetition), where the target forms are real words or nonwords containing an onset cluster. Like in Den Ouden (2002), the functional state of the speech production mechanism is deduced from the combination of performance results on the different tasks. It is found that children perform poorly on the PN task in the initial sessions, while they do better on the NWR and/or the WR tasks. This points to the lexical representation as the initial error locus because performing successfully on the NWR and/or the WR task does not require lexical access. In later sessions, the error pattern changes. Like in Chapter 3, these changing error patterns are taken to reveal developments in the speech production mechanism, and they are discussed in detail.

In **Chapter 5**, I turn to perception, and ask how detailed the representation of onset clusters is in the child's mental lexicon. Do children exhibit different looking behavior when they perceive correctly produced onset clusters as opposed to reduced onset clusters? If this is the case, the segmental representation can be assumed to be detailed, containing both C_1 and C_2. If not, omissions in production could be the result of incomplete segmental representations. I examine two-year-olds' perception of correct vs. reduced /sC/ clusters, like in the word *stoel* /stul/ 'chair' and /C+liq/ clusters like in the words *trein* /tʀɛin/ 'train' and *bloem* /blum/ 'flower'. Interpreting the looking times in line with earlier work on children's perception of mispronounced words (Swingley & Aslin, 2000, White & Morgan, 2008), results seem to indicate that two-year-olds exhibit awareness of /sC/ cluster reduction but not of /C+liq/ cluster reduction. However, another interpretation of the results is that the longer looking times actually indicate that the correct form is novel to the child, and therefore attracts longer attention. This interpretation is strengthened by the children's performance on a small production task, where they simply had to name the pictures that were shown in the perception experiment, and where we find that in children who have not acquired /sC/ clusters yet, this novelty effect is stronger than in children who have already acquired /sC/ clusters.

Finally, in **Chapter 6**, the results obtained in Chapters 2 to 5 are discussed in relation to each other, and I will summarize what the combination of results can tell us about the developing speech production mechanism.

Chapter 2. From toddlers' mouths to adults' ears: production and perception of reduced onset clusters occurs on different channels[1]

2.1. Introduction

Phonetic or phonological cluster reduction is a common phenomenon in young children's speech productions. In this chapter cases of cluster reduction in word onsets are discussed in which the child apparently omits the second consonant, as in [dʌk] for *truck* and [siːp] for *sleep*. The research discussed here addresses two questions. First, we want to find out whether toddlers intend to express a complex onset despite the apparent omission of the second consonant. Does the lexical representation of a reduced cluster contain information about the omitted consonant or not? 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 purpose of performing a detailed analysis of the reduced form is to help to determine the source of the deviation from the adult target form. Our acoustic analyses indeed reveal traces of the omitted consonant. This leads to our second question, namely whether adults can distinguish children's words with reduced onsets from words starting with an identical simple onset when these are presented next to each other. In other words, when adults are asked to pick from a child's minimal pair the production that has an onset cluster in the adult language, can they use the acoustic trace of the "omitted" consonant as a reliable cue? Here we find that adult listeners use different cues for their decisions than the cues that the child provides.

[1] This Chapter is identical to the manuscript: Gulian, M, Levelt, C. & Boersma, P. *From toddlers' mouths to adults' ears: production and perception of reduced onset clusters occurs on different channels*. It therefore uses the first person plural instead of singular. The manuscrpt is ready for submission to a linguistic journal.

2.1.1. Theoretical background

One of the goals of the present study is to get better insight in the way consonant clusters are stored and handled in the toddler's mental lexicon and speech production mechanism. Do toddlers store adult cluster words as CV- (consonant–vowel) sequences or as CCV- sequences underlyingly, and if this is the case, where in the production process does the reduction take place?

To explore the possibilities, we suggest the heuristic model of speech production in Figure 1, which combines phonological and psycholinguistic views of the levels of representation involved (Levelt et al., 1999; Boersma, 2011). In 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 phonological underlying form, which contains the stored information about the word's sounds. From this information a phonological surface form is created in the phonological production process. Subsequently, phonetic implementation may convert this surface form to an auditory-phonetic target (for adults: MacNeilage, 1981, Gay et al., 1981; for children: Oller & MacNeilage, 1983), which is then translated by sensorimotor knowledge to an articulatory-motor program that controls the speech muscles. The precise steps in the whole process are subject to debate, but Figure 1 will help us formulate hypotheses about the localization and causes of reduction.

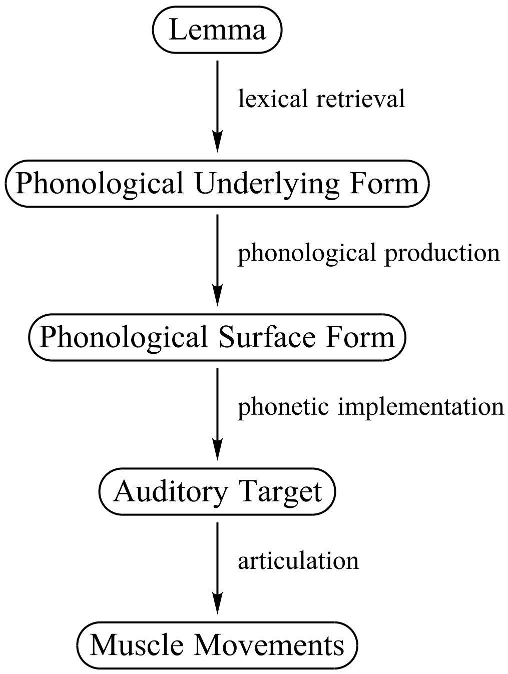


Figure 1: Heuristic speech production mechanism

Figure 1 suggests at least seven potential locations or causes for cluster reduction.

The acoustic signal will have different characteristics depending on the locus of reduction. Consider the Dutch adult word pair [bʀoːt] 'bread' versus [boːt] 'boat', and assume that the child stores 'boat' as /boːt/ in her underlying form. The question now is: where does the child reduce the adult's /bʀ/ in [bʀoːt] ('bread') If the child's underlying form for 'bread' is /boːt/, identical to the one for 'boat', then the child appears to have reduced the cluster either (1) already somewhere in her comprehension of the adult word, or (2) when storing the word in her lexicon for the first time, perhaps as a result of a morpheme-structure constraint; in these cases, we predict that the child will pronounce 'bread' in an identical way to 'boat' at the acoustic level. If the child's underlying form for 'bread' is /bʀoːt/, but her surface form is /boːt/, then either (3) her phonological grammar dictates that underlying /bʀ/ should correspond to a surface /b/, or (4) the surface form is restricted by a structural constraint such as */CC/; in these cases, the reduction is again discrete (i.e. all or none), so that complete acoustic homophony with the production of 'boat' is predicted. If the child's underlying and surface forms are /bʀoːt/, it is possible that (5) she has trouble mapping the surface /bʀ/ to the appropriate auditory cues, thus targeting something close to, but not necessarily identical to, [boːt]; in this case, the reduction is not discrete at the acoustic level, but a transcriber may classify the sound as the phonological surface form /boːt/ with her adult Dutch perception system. In this case we predict that the child may object to an adult pronouncing 'bread' as [boːt] (i.e. the *fis* phenomenon: Berko & Brown, 1960). If the auditory target is a full-fledged [bʀoːt], the articulatory result may still be close to [boːt] as a result of (6) a sensorimotor mapping that does not yet link the auditory cues with the appropriate muscle gestures (Ferguson & Macken, 1983) or (7) developmental restrictions on the planning or timing of muscle gestures (Studdert-Kennedy, 1987); in these cases we may find an acoustic trace of /ʀ/, although a Dutch transcriber might not notice this. Therefore, if we analyze the child's acoustic productions of 'bread' and do find a trace, then we can conclude that reduction has taken place by one of the mechanisms (4) through (7); if there is no trace at all, the cause may lie in mechanisms (1) through (3).

Gradient versions of these mechanisms are also possible. It could be the case, for instance, that (due to a comprehension restriction, a lexical restriction, or a surface restriction) the child's surface structure is the reduced segment sequence /CV/ but does exhibit in the vowel an extra feature, for instance rhoticity, that somehow expresses the reduced C2. Thus, 'bread' could be represented as /bo^{+rho}ːt/. The extra feature would typically come with fewer auditory cues for the adult listener than a segment would, so that an intended /bo^{+rho}ːt/ will be perceived by an adult listener as a complete homonym of /boːt/

'boat'. If this is the case, an acoustic trace of /ʀ/ may be found in the child's realization of 'bread'.

2.1.2. Covert contrasts in the literature

Studying the acoustic waveforms of toddlers' productions is an interesting way to find out more about the lexical representations of early words. Up until now, young children's lexical representations have mostly been studied using perception experiments (e.g. Fennell & Werker, 2003; Swingley, 2003; Swingley & Aslin, 2000, 2007; White & Morgan, 2008; for an overview see Newman, 2008). However, a detailed analysis of children's productions gives a different perspective on the issue, and directly confronts the difference that exists between detailed representations and reduced productions (Pater & Barlow, 2003; Smolensky, 1996).

Acoustic analyses have led to the discovery of a number of "covert contrasts" in toddler's productions (for an early overview see Scobbie, 1998). McLeod et al. (1998) showed that Australian English two-and-a-half-year-olds pronounce a [k] reflecting a target /sk/ cluster with a shorter VOT than a [k] reflecting a target singleton /k/ onset. Carter and Gerken (2004) analyzed truncations in two-year old children who had to repeat sentences like *He kissed Lucinda* – *Lucinda* being a ready target for reduction in toddler speech – and *He kissed Cindy* and found a larger time gap between *kissed* and reduced *cinda* than between *kissed* and correct *Cindy*. Song and Demuth (2008) recorded longitudinally three children (1;6 – 2;6) and found in their utterances differences between reduced target coda clusters and similar correctly produced target singleton forms: compensatory vowel lengthening was found in case the coda cluster was reduced. Lowenstein and Nittrouer (2008) showed that American-English two-year-olds produce voiceless target plosives with longer VOTs than voiced target plosives, although the two transcribers could not perceive this difference. Gulian and Levelt (2011) found that Dutch two-year-olds pronounced reduced article-noun phrases with a reduced cluster differently from singleton counterparts.

The authors compared phrases like *een peen*, where *peen* [peːn] was the reduced form of *speen* (/speːn/ 'pacifier') with *een peek*, where *peek* [peːk] was the intended singleton nonword *peek* /peːk/. They found that there was a larger time interval between the nasal in *een* and the plosive in *peen* as compared to the same interval in *een peek*.

All of these studies thus reveal knowledge that language learners have, but do not make audible in a way that adult listeners can perceive.

In the two studies below, we focus on two clusters that are very often reduced in Dutch child language productions, namely /Cr/ (plosive + rhotic[2]) and /kn/. In study 1, word productions with reduced renditions of these target clusters are analyzed acoustically and compared to productions of corresponding words with singleton onsets. Thus, an adult onset cluster /Cr/, apparently produced by the toddler as [C-] is compared to the toddler's production of a phonetically similar word with an adult singleton onset /C-/. For instance, the utterance [boːt] for *brood* 'bread' is compared to *boot* [boːt] 'boat'. An example of the other cluster type is *knippen* (adult target [knɪpə]) 'to cut', produced by the child as [kɪpə], which is compared to *kippen* [kɪpə] 'chickens'. In study 2 we test the way adults perceive these minimal pairs in toddler speech.

2.2. Study 1: Child production of /Cr/~/C/ and /kn/~/k/ word pairs

In order to answer the question where in the production model cluster reduction originates, we concentrate on /kn/ and /Cr/ cluster types in Dutch. Specifically, we look for the productions of minimal pairs of singleton and cluster targets, e.g. for cases in which the same child produced both 'bread' (adult target [bʀoːt]) and 'boat' (adult target [boːt]), or for cases in which the same child produced 'chickens' (adult target [kɪpə]) as well as 'to cut' (adult target [knɪpə]).

[2] In this position, Dutch has only one rhotic phoneme, which can be realized as [ʀ], [r] or [ɾ] (Sebregts, 2015).

Any small systematic acoustic difference between the members of a produced pair could indicate that the child intends to make a difference between the word forms, even if both members were transcribed identically by adult researchers (namely with a single consonant).

2.2.1. Participants

For this study we looked for young Dutch monolingual children who reduced /Cr/ and/or /kn/ cluster words in their speech. We found such children in two separate datasets, namely in the existing CLPF database (Levelt, 1994; Fikkert, 1994), available in Childes/Phonbank (Rose & MacWhinney, 2014), and in our own new recordings at day-care centers collected specifically for the present purpose. The CLPF database consists of longitudinal recordings of 12 children acquiring Dutch as their first language, aged roughly between one year and two and a half years at the start of the data-collecting period. Currently, audio files are available for 6 of the 12 children, and in the data of 4 of these children we could find the necessary word-pairs for comparison. The day-care-center dataset was collected by the author of this thesis by recording 30 toddlers with a mean age of 2;1 years at four Dutch day-care centers. Nine of these children already produced /kn/ and /Cr/ target clusters in an adult-like manner and were excluded from the analyses. The data for analysis thus included the productions of four children from the CLPF database and 21 children from the day-care-center recordings, forming a total of 25 children. Eight of these 25 children reduced both /Cr/ and /kn/ clusters, while the remaining 17 children reduced either one cluster or the other (see Appendix 2).

2.2.2. Method: /Cr/~/C/ word pairs

We start out by discussing the acoustic analysis of target /Cr/~/C/ word pairs.

2.2.2.1. Participant selection

Here we consider data from those four children from the CLPF database whose speech exhibited the phenomenon of /Cr/ cluster reduction and who also produced a singleton counterpart in the same session or in a closely related session, and data from those 11 day-care-center children who reduced the /Cr/ cluster at least once and who produced at least one singleton counterpart. The mean age of the 4 database children at the times of the recordings that are used here was 2;1.6 (age range 1;8.10 - 2;4.26). The mean age of the 11 day-care-center children was 2;0.29 (age range 1;6.0 - 3;0.1).

2.2.2.2. Data selection

In order to detect acoustic traces of reduced /Cr/ clusters, word pairs were compared for each child separately. For every child, we paired a word production with a reduced target cluster with a word production with a singleton onset consonant that matched the onset consonant of the reduced form as closely as possible.

At the day-care centers, toddlers were asked to repeat a list of Dutch words with initial clusters and matching words (or sometimes non-words) with a simple onset (i.e. *trein* 'train' matched to *Thijs* (a common Dutch boy's name). If possible, pictures of the words-to-be-repeated were used to encourage production. A list of the words that the children had to repeat is given in Appendix 1. The children's utterances were recorded with 16-bit 44100-Hz sampling with a Microtrack II digital recorder and an external Microtrack II microphone. At the time of each recording, the responses by the 30 children were also transcribed online. Later, the online transcriptions containing cluster reductions were selected for more detailed off-line phonetic transcriptions, which I checked first, and were subsecuently checked by an experienced phonologist.

Word productions were determined to be reduced cluster words if in the data the target word contained an onset cluster and according to the phonetic transcription of the word produced by the child, the second consonant of the consonant cluster was omitted[3], such as in the transcription [boːt] for intended *brood* 'bread' [bʀoːt]. In the CLPF database a search was carried out to find all utterances that contained a cluster in its adult target form but missed the second consonant of that cluster in the phonetic transcription of the child's actual production. The matching singleton-onset word would ideally form a minimal pair with the target cluster-onset word, differing only in the absence of a second consonant. For *brood* [bʀoːt], for example, the ideal match was *boot* 'boat' [boːt].

From now on the target singleton-onset words such as *tijd* and *boot* will be referred to as /C/ words. If no ideal match could be found for a /Cr/ word, a /C/ word was selected that shared as many features as possible with the onset plosive and the subsequent vowel, e.g. the /Cr/ word *trein* 'train' /tʀɛin/, produced as [tɛin], was paired with *Thijs* (Dutch boy's name) /tɛis/ in the analysis. If the utterances selected for analysis were polysyllabic, such as *draaimolen* 'merry-go-round' /ˈdʀaːiˌmoːlə/, they were always stressed on the first syllable. The /Cr/ word and the matching /C/ word were always produced by the same child and originated from the same recording session, or from recording sessions that were no more than 1 month apart.

After the strict selection criteria for matching word pairs, in the end the analysis of the /Cr/ clusters is based on 47 word pairs, i.e. 47 target /Cr/ words matched to 47 /C/ words. Of these 47 word pairs, 21 came from the CLPF database, and 26 from the day-care center set. A list of the reduced cluster words and their matching singleton consonant forms is given in Appendix 3.

---

[3] The authors of the CLPF database, C. Levelt and P. Fikkert, were the primary transcribers of the CLPF dataset, and I made the transciption of the day-care-center dataset.

2.2.2.3. Measurement method

In this study we looked at words with target onset /Cr/ clusters where the second consonant was apparently omitted, and compared them acoustically to similar /C/ words. All the acoustic measurements presented in this chapter were made using Praat 5.0.10 (Boersma & Weenink, 2008).

To minimize the chance of subjective measures, the assistant who carried out the acoustic measurements was blind to the actual transcriptions of the words produced by the children. All 94 utterances (which consisted of the 47 reduced cluster words and 47 matching simple onset words) were anonymized[4]. If necessary, the utterances were trimmed back to only the initial consonant-vowel sequence; for example in the utterances [boːt] (from *brood* 'bread') and [boːtə] (from *boten* 'boats') the final part was removed, leading to [boː] and [boː] respectively, to prevent them from revealing the word meaning to the assistant, who was told to determine the vowel and utterance boundaries as if the word was a /C/ word.

Two formants, F2 and F3, were measured at two points in time in the spectrum by means of a band-filter analysis method that had been used before for the description of infant vowel productions (Wempe, 2001; Van der Stelt et al., 2005). This method, which takes into account the child's pitch to estimate a spectral envelope representation of an utterance, has the advantage above LPC (linear predictive coding, the most widely used formant analysis method) of being less sensitive to the incorrectly chosen parameters that are likely to occur with LPC if there are high pitches in the data, which is often the case in child speech.

The approximant rhotics [ɹ] and [ɻ], such as occur in English and in Dutch codas, tend to come with a low F3 and to some extent a low F2 (Lindau, 1985; Plug & Ogden, 2003; Scobbie & Sebregts, 2011). The adult realization in Dutch /Cr/

[4] After all analyses were carried out, the anonymous utterances were linked to the original transcriptions again.

clusters is more likely the uvular trill [ʀ] or the alveolar trill or tap [ʀ, ɾ]. Gulian (in prep.) found modest F3 and F2 values for these variants: in going from the rhotic to the vowel there was a rise in F2 and F3 for front vowels and a lowering in F2 for back vowels, both for the alveolar and for the uvular variants, with the alveolar trill exhibiting a slightly lower F3 than the uvular trill. If we detect similar formant movements in the vowel onset in reduced cluster words in toddler speech, this could therefore be considered a likely trace of the omitted rhotic.

In order to be able to measure F2 and F3 movement in the vowel, the two formants were first measured at the immediate vowel onset (t_1), followed by a measurement at one quarter of the entire duration of the vowel (t_2). Figure 2 gives an example of a child production of the word *kraan* 'faucet' /kʀaːn/, where the target rhotic is fully produced and the raising of F2 and F3 can be clearly discerned.

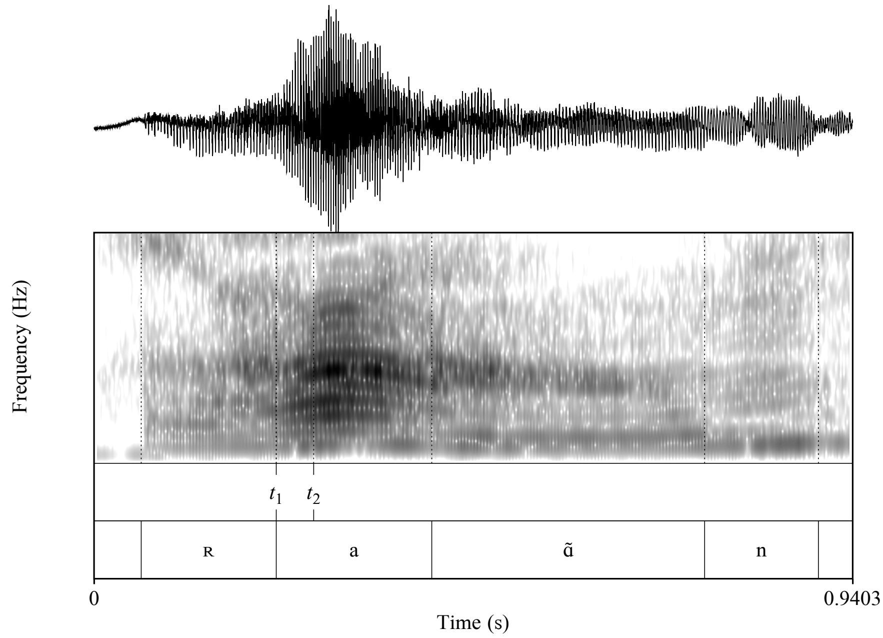


Figure 2: Waveform and spectrogram of the word *kraan*, produced [ʀaãn] by a child aged 3;3 (t_1 and t_2 are the time points where formants are measured).

In order to capture a possible upward movement in the vowel onset, the values for F2 and F3 at times t_1 and t_2 are used for a simple calculation: the Hz value at t_1 is subtracted from the Hz value at t_2. When the obtained value is positive, the formants have moved upward, and this is interpreted as indicating the intended presence of a preceding rhotic. For example, in Figure 2, the calculations for F2 and F3 both result in positive values (a difference of respectively 268 Hz and 24 Hz). This confirms that both F2 and F3 are rising in the vowel onset, although /a/ is not a prototypical front vowel. If the value is negative, the formant in question appears to have lowered in the vowel onset, which is interpreted as indicating a traceless omission of the target rhotic in the production.

In addition to F2 and F3 movement in vowel onsets, we measured vowel and utterance duration. Since we were interested in acoustic traces of the second consonant in reduced onset clusters, we had to be on the look-out for forms of compensatory lengthening. For the utterance duration measure, the duration of the vowel plus preceding and following coda consonants (in the cases where the coda was not trimmed) was measured.

To summarize, for the /Cr/~/C/ minimal pairs, such as the words *trein* and *Thijs*, where *trein* is produced as [tɛin], four different measures were taken. All 47 pairs (94 utterances) were measured for their vowel duration, their utterance duration, and the F2 and F3 movement in the vowel onset. In the result section (2.4) we turn to the outcomes of these four acoustic measures and determine whether any of these measures was distinctive for the minimal pairs. Below we first discuss the different measures we took with the other cluster type in this study.

2.2.3. Method: /kn/~/k/ word pairs

This section discusses the analysis of the reduction of target /kn/~/k/ word pairs.

2.2.3.1. Participant selection

Utterances from both the CLPF database and the day-care-center recordings are analyzed. We use utterances from two children from the CLPF database, with a mean age of 2;3 (age range 2;0 – 2;6)[5]. These children overlap with the children that exhibited /Cr/ cluster reductions in this study (Study 1, see Appendix 1). From the 30 children who were recorded at day-care centers, 16 children (mean age 1;11, range 1;8 - 3;0) reduced the /kn/ clusters. In total we analyze the /kn/~/k/ utterances of 18 children.

[5] As in the /Cr/~/C/ acoustic analysis, recordings from different sessions from the same children were used.

2.2.3.2. Data selection

In order to detect acoustic traces of reduced /kn/ clusters for each child, again pairs of words were compared. As in the previous study, a reduced onset-cluster utterance was selected and compared to a singleton-onset utterance that was closely matched to the reduced cluster production. The selection criteria for the matching word are identical to those mentioned in 2.1. Here /kn/ words, such as *knippen* 'to cut' and *knoop* 'button', which according to their transcriptions were produced as [kɪpə] and [ko˙p], were compared to words like *kip* 'chicken' [kɪp] and *kopen* 'to buy' [koːpə].

For the present analysis, 37 target cluster words were matched to 37 target words starting with a singleton /k/. From these 37 pairs, six were produced by the two children from the CLPF database, while the remaining 31 pairs were produced by the children from the day-care center. A list of the /kn/~/k/ word pairs is given in Appendix 4.

2.2.3.3. Measurement method

As in the /Cr/ study, the assistant who carried out the acoustic analyses was blind to the transcriptions and to the intended form of the utterances. Three duration measures were taken: vowel duration, utterance duration, and an additional duration measure called "vowel complexity". The vowel complexity measure was first suggested in a pilot study (Gulian &| Levelt, 2008), which observed that the vowel onset in reduced /kn/ words often exhibited an atypical diphthongization. An example of an utterance containing a diphthongized vowel onset is given in Figure 3. In order to determine the vowel complexity measure we use the following criteria: if the first half of the vowel exhibits a changing vocalic pattern, i.e. a diphthongization, then the duration of the first part (labeled as "$part_1$" in Figure 3) constitutes the vowel complexity value for this item; otherwise, if the first half of the vowel does not exhibit a change, the vowel complexity value is taken to be zero. In case of doubt, the vowel complexity measure is taken to be zero as well. In order to become acquainted with the determination of this duration measure the assistant was familiarized with productions from the pilot study, containing similar diphthongized vowel onsets, which were not part of the list of 74 stimuli from the present data set.

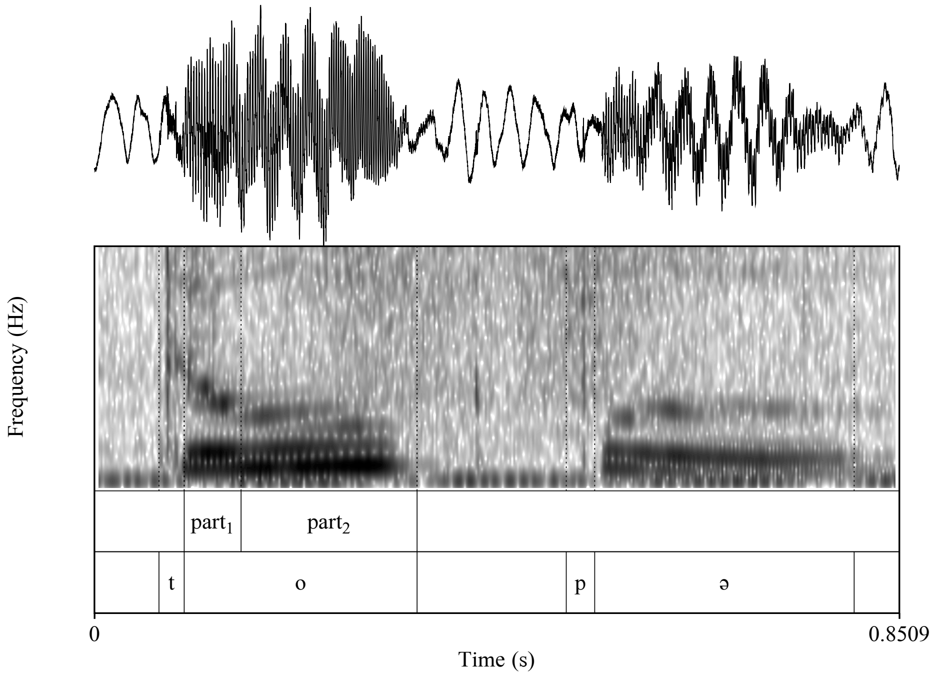


Figure 3: Waveform and spectrogram of the word *knopen* (actual production [topə]) by a child aged 2;5. $Part_1$ and $part_2$ stand for the two parts within the vowel that are revealed in the spectrogram.

In addition to the three duration measures, we took three other measures that could point to traces of the omitted nasal in the vowel. In order to do this we selected three measures from a list of acoustic characteristics of nasalized vowels described in Pruthi et al. (2007). According to Pruthi et al. (2007), nasalized vowels exhibit, among others, reduction in the first formant (F1) amplitude[6] and in the overall intensity of the vowel, and a movement of the low-frequency center of gravity towards a neutral vowel configuration (besides these three, Pruthi et al. studied another four acoustic correlates of nasality, but we chose not to measure them because they were all related to the F1 measure).

The three measures we took (mean F1, overall vowel intensity and mean center of gravity of the vowel following the plosive) were analyzed by means of a script in *Praat*. To measure mean F1, the settings for formant measurement in *Praat* were adapted to toddlers' voice quality, namely to search for up to 5 formants in the range from 0 to 6,000 Hz. All measures were carried out for both reduced /kn/ words and singleton /k/ words.

To sum up, for the reduced nasal clusters in Dutch we measured six acoustic characteristics, which consisted of comparing the word pairs in terms of utterance duration, vowel duration, "vowel complexity duration", mean F1, overall vowel intensity and center of gravity.

2.2.4. Results of Study 1

2.2.4.1. Results: /Cr/~/C/ word pairs

In the case of /Cr/~/C/ word pairs, we carried out four acoustic measures: vowel duration, word duration, F2 and F3 movement in the vowel onset. Our question was whether the target complex onset words would differ from target simple onset words in any of these four acoustic measures, even though they would all be produced with a simple onset. For this purpose we conduct a repeated-measures multivariate analysis of variance with word type (reduced cluster vs. simple onset) as repeated factor and four acoustic measure types (vowel duration, utterance duration, F2 movement and F3 movement) as dependent variables.

[6] The reduction of F1 amplitude is especially true for low vowels (Pruthi, 2007; p. 3871), while for high vowels, nasality brings F1 higher.

The analysis of variance reveals that word type has a significant effect on one of the acoustic measures, namely F2 movement (F [1,46] = 4.97, p = .031). This significant effect shows that F2 movement tends to be positive (thus upwards) in the vowel onset of reduced /Cr/ cluster words as compared to simple onset words (M = 81 Hz, SD = 298 Hz, and M = -43 Hz, SD = 282 Hz, respectively); see Figure 4. Word type did not show a significant effect on the other dependent variables (for all three: $p \geq .359$).

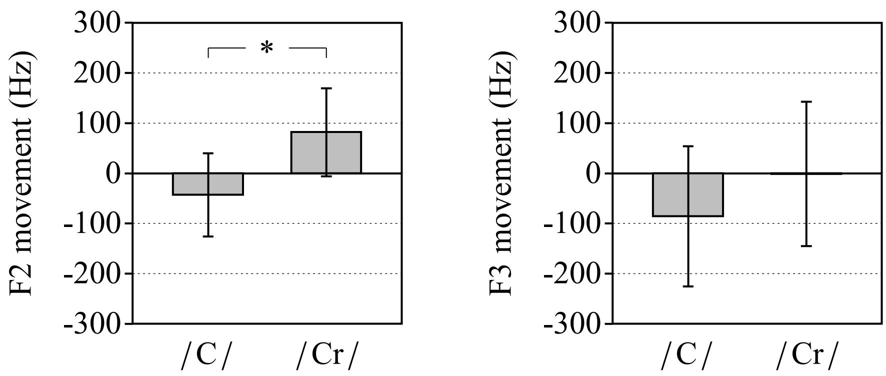


Figure 4: F2 and F3 movements in the vowels of target /C/~/Cr/ word pairs, showing means and 95% confidence intervals.

2.2.4.2. Results: /kn/~/k/ word pairs

The /kn/~/k/ word pairs are compared using six different acoustic measures: vowel duration, word duration, vowel complexity duration[7], center of gravity, overall intensity and mean F1. Because the type of the following vowel determines the height of F1 of the nasal (when present), we take into account

[7] Section 2.3.3 gives a clarification of this unconventional measure.

whether the vowel in the word pairs is high or not. From now on we refer to these two word types as /kni/ and /ki/ words versus /kna/ and /ka/ words.

A repeated-measures MANOVA with word type (reduced cluster vs. simple onset) as the repeated factor, vowel type (onset + /i/ vs. onset + /a/) as the between-items factor, and six dependent variables (vowel duration, word duration, vowel complexity duration, center of gravity, intensity and F1) reveals a substantial main effect of vowel type (F [6,30] = 6.633, p < .001) and a main effect of word type (F [6,30] = 2.336, p = .057). We find no interaction between vowel type and word type (all $p \geq$ .164). The univariate ANOVAs reveal an effect of vowel type on one of the six acoustic measures: mean F1 (F [1,35] = 32.8, p < .001). If we consult the descriptive statistics for this factor, we conclude that /ki/~/kni/ words tend to have lower F1 than /kna/~/ka/ words (see Table 1 and Figures 5 and 6).

Table 1: Descriptive statistics of the acoustic measures that show a significant interaction either with vowel type (the first pair) or with word type (the last two pairs).

Measure and word type	**Mean (SD)**
F1: /ki/, /kni/ words vs. /ka/, /kna/ words	554 (105) Hz vs. 735 (130) Hz
vowel complexity: /kn/ words vs. /k/ words	0.018 (0.025) s vs. 0.006 (0.019) s
center of gravity: /kn/ words vs. /k/ words	656 (267) Hz vs. 760 (323) Hz

Another statistically significant effect that is found is the effect of word type on vowel complexity: F [1,35] = 5.884, p = .021, and a nearly significant effect on center of gravity: F [1,35] = 3.414, p = .073, see Table 1 and Figures 5 and 6. Word type did not show an effect on the other dependent variables (all $p \geq$ .405).

As the mean values of these measures show, the vowel complexity duration tends to be longer in reduced /kn/ words than in simple onset /k/ words, while center of gravity tends to be lower for reduced clusters than for simple onsets.

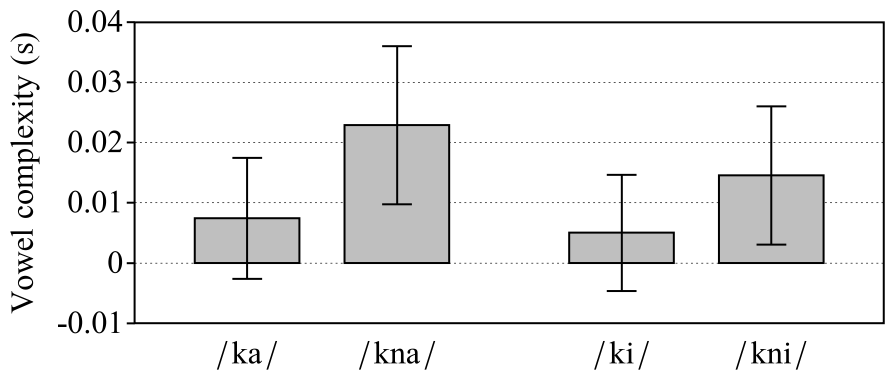


Figure 5: Vowel complexity measure in low (left) and high (right) vowels of target /k/~/kn/ word pairs. The values on the *y*-axis are presented in seconds.

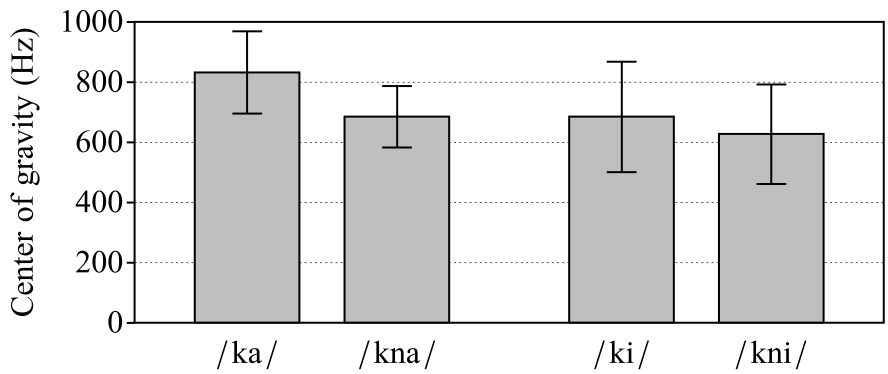


Figure 6: Center of gravity measure in low (left) and high (right) vowels of target /k/~/kn/ word pairs. The values on the *y*-axis are presented in Hz.

2.2.4.3. Summary of the results

Regarding the acoustic measures, the following can be concluded. For reduced /Cr/ words, the acoustic characteristic that seems to distinguish them from simple onset /C/ words is F2 movement in the vowel onset, showing a rise in words with reduced clusters. Word, vowel duration and F3 movement were not distinctive between the two sets of words (see Figure 4). As for reduced /kn/ words, vowel complexity duration was longer for the reduced cluster words than for the simple onset words and center of gravity showed a trend of being lower for the reduced /kn/ words (see Figures 5, 6). None of the other acoustic measures of nasality were distinctive. For both types of cluster reduction, then, Dutch toddlers produce some of the acoustic characteristics of the target second consonant of the cluster.

Given the fact that acoustic traces of the second consonant of both target /Cr/ and /kn/ words can be found in the productions of Dutch toddlers, the next question is whether or not adult listeners are able to pick up on these subtle cues. Since the productions of target words with onset clusters that were used for the acoustic measurements had been transcribed – by trained linguists – with singleton onsets, this does not seem likely in a natural context. However, would it be possible to find evidence for listeners picking up on the cues in an experimental setting? This question was addressed in Study 2.

2.3. Study 2: Adult perception of reduced target clusters /Cr/ and /kn/

In this study Dutch adult listeners participated in a forced-choice identification task, where the word-pair stimuli used in the first study were presented without the situational and word context that could help to disambiguate the two utterances. We asked participants to try to identify which of the two words that formed a minimal pair was originally an onset cluster word. The question was whether adult listeners would be able to rely on the acoustic cues provided by the child to correctly identify the reduced-cluster [CV] sequence from its singleton counterpart [CV] sequence.

2.3.1. Method

2.3.1.1. Stimuli: Word pairs with onset clusters /Cr/ and /kn/

The stimuli were the same CV sequences as those in Study 1. See the description of the stimuli in 2.2.2.2 and 2.2.3.2.

2.3.1.2. Procedure

The forced-choice identification task was carried out using the computer program *Praat*. Before starting, the test participants received information about how the minimal pairs had been obtained and why they heard only the first CV(C) sequence[8] of the child utterances. They were told in advance that perceptually the two utterances hardly differed from one another and that a consonant cluster was not perceivable, but that nevertheless in each test trial one of the two utterances was a target onset-cluster word. Finally, the adults were instructed that their task was to try their best to identify or, if necessary, guess which of the two stimuli corresponded best to a target word with an onset cluster.

Each participant was seated in front of a computer where he or she saw a gray screen with the instruction in the upper part of the screen saying "Choose the word that seems to start with a consonant cluster" and the words "first" and "second" that appeared on two yellow buttons in the screen center (all written in Dutch). The participants were not told what the possible target words were. In other words, when they heard [bo:], [bo:], they were not told to choose between the words *brood* and *boot*. Each participant simply heard two stimuli in a row and had to click on the left button ("first") when the participant thought the first of the two utterances was more likely to be the cluster word and on the right button ("second") when the participant thought the second utterance was more

[8] From now on we refer to all trimmed stimuli as words.

likely to be the cluster word. The participants heard the stimuli through a Sennheiser headphone set and the stimuli were presented to them only once. After hearing both utterances they were “forced” to make their decision.

The first three trials familiarized the participants with the test procedure. The experiment itself consisted of two parts, one for /Cr/~/C/ minimal pairs and one for /kn/~/k/ minimal pairs. The order of the two parts of the experiment was randomly distributed among the participants, so that 17 of them started with the /Cr/~/C/ part of the test and 18 of them started with the /kn/~/k/ part of the test. The /Cr/~/C/ part consisted of 47 trials and the /Cr/~/C/ part of the experiment consisted of 36 trials. In total, the experiment lasted about 10 minutes.

2.3.1.3. Participants

In total 35 native speakers of Dutch (19 women, 16 men) performed the forced-choice identification task. Mean age was 35.8 years (age range 23–60). Of these, 12 participants were exposed to the speech of toddlers regularly.

2.3.1.4. Analysis

Each response in the forced-choice perception task was labeled as correct or incorrect depending on whether the participant choice matched the toddler’s intention. For instance, if in the ‘bread’ ~ ‘boat’ pair the participant chose the intended ‘bread’ word as the cluster word, this choice was deemed correct. Pooling the results over all 35 listeners, we obtained a ‘correct’ score between 0 and 35 for each item pair.

As we were interested in the factors that influence the choice of adult Dutch listeners when they are forced to decide between a cluster and a singleton utterance produced by a toddler (although these two utterances sound almost the same), we submitted the correctness scores from the forced-choice identification test to a logistic-regression analysis. The predictors (factors) in this logistic regression are the differences between the acoustic measures of the cluster word and the acoustic measures of the singleton word. For instance, for the /Cr/~/C/-word set, Study 1 measured four acoustic cues (F2 movement in the vowel onset, F3 movement in the vowel onset, vowel duration and word duration) for both members of each of the 47 utterance pairs, so that the logistic regression for the /Cr/~/C/ word pairs involves the following four factors: F2 movement difference (the F2 movement of the cluster word minus the F2 movement of the corresponding singleton word), F3 movement difference, vowel duration difference, and word duration difference. This results in four difference measures for each of the 47 pairs. Likewise, the /kn/~/k/ word set leads to six difference measures for each of the 37 utterance pairs.

The interpretation of the coefficients that result from the regression analysis is as follows. If, for instance, the estimated coefficient for the vowel duration difference is positive, this would indicate that listeners have a greater chance of scoring correct if the cluster word is longer than the singleton word than if the cluster word is shorter than the singleton word; we could then conclude that listeners associate cluster words with longer duration and singleton words with shorter duration.

Following the same line of reasoning, if the estimated coefficient for the vowel duration difference is negative, this would indicate that listeners have a greater chance of scoring correct if the cluster word is shorter than the singleton word than if the cluster word is longer than the singleton word. If this would be the case, then we could conclude that listeners associate cluster words with a shorter duration and singleton words with longer duration.

2.3.2. Results

2.3.2.1. /Cr/~/C/ word pairs

Here we discuss the overall scores in the identification test, which are calculated in percentages. As explained in the previous section, /Cr/~/C/ word pairs were presented to the listeners and their task was to identify the target /Cr/-word from the two. The answers were identified as being either correct or incorrect. Higher 'correct' scores indicate a higher sensitivity of the participant to the acoustic cues that distinguish the child's /Cr/ production from her /C/ production. The results are sorted by item and are used in the logistic regression analysis where the acoustic predictors for correct perception are investigated.

In the identification task, the mean result for correct identification per item was 51%, the range being 2.8% to 91.4% (N = 47). The large variation results from the fact that some reduced cluster words were easy to identify, while others were incorrectly taken for simple onset words. Our next question then was: are the correct identifications, although at chance level, correlated with any of the acoustic cues that were found to be distinctive in Study 1? In order to find an answer to this question we obtained a difference value for the four cues from Study 1 (F2, F3, utterance duration and vowel duration) for each of the utterance pairs from the forced-choice identification task, by subtracting the value obtained for the simple onset word from the value obtained for the reduced cluster word.

In order to find out which of the four cues are good predictors for the correct judgment of a certain word pair, we conduct a logistic regression analysis (with the function *glmer* in *R*). The response variable was whether the participant gave a correct or incorrect answer, where an answer was considered correct if the participant chose the intended cluster word by the toddler as the cluster word. The potential fixed factors in the logistic regression were the F2 movement difference (dMovF2), the F3 movement difference (dMovF3), the vowel duration difference (dDurV) and the word duration difference (dDurW). Finally, the participants are included in the regression model as a random factor.

With only participant as a factor, the intercept of the model becomes +0.04038. This means that the average log odds in favor of a correct choice is +0.04038, i.e. a correct choice is exp(+0.04038) = 1.0412 more likely than an incorrect choice. This corresponds to the average percentage correct mentioned above (51%/49% = 1.04). The standard error of the log odds is 0.06940, so the 95% confidence interval of the log odds runs from -0.09567 to +0.17643 (i.e. the odds correct could be anywhere between 0.909 and 1.193), which includes zero, so we have not shown that listeners can differentiate the average pair better than chance.

Of the four fixed factors, entering dDurW yields the best improvement of the model ($p = 1.18 \cdot 10^{-8}$), although entering dDurV would have yielded a substantial improvement ($p = 5.4 \cdot 10^{-4}$) as well. After adding dDurW, adding ddF2 into the model yields another improvement, and after this no additions, not even of dDurV, improve the model any further.[9]

In the final model, dDurW has a coefficient of 1.4368345/s (standard error 0.2463989/s; c.i. = +0.95380 .. +1.91987/s; p from zero = $5.5 \cdot 10^{-9}$). This means that every increase of 10 ms in the duration difference between cluster word and singleton word increases the odds for a correct score by a factor of exp(0.014368345) = 1.0145, i.e. by 1.45% (c.i. = +0.96% .. +1.94%). This means that the longer the cluster word lasts, the more often it is correctly chosen to be the cluster word, and the longer the singleton word lasts, the more often this is incorrectly chosen to be the cluster word. The unsurprising interpretation is that

---

[9] The cause that dDurV cannot be included in the model, although it had such a good p value when entering it on its own, is that dDurV and dDurW are strongly correlated, so that entering dDurW takes away dDurV's chances.

listeners associate cluster words with a longer duration and singleton words with a shorter duration.

The other factor in the final model, dMovF2, has a coefficient of -0.0002874/Hz (standard error 0.0001350/Hz; c.i. = -0.00055205. -0.00002275/Hz; *p* from zero = 0.0332). This means that every *decrease* of 10 Hz in the F2 movement difference between cluster word and singleton word increases the odds for a correct score by 0.287% (c.i. = +0.023% .. +0.554%). This means that the more F2 falls (or the less it rises) in the cluster word, the more often it is correctly chosen to be the cluster word, and the more F2 rises (or the less it falls) in the singleton word, the more often this is incorrectly chosen to be the cluster word. The interpretation is that listeners associate cluster words with a falling F2 and singleton words with a rising (or less falling) F2.

It thus appears that word duration[10] is the main successful predictor for the participants' judgments. Given that the correlation found is positive, it means that correct results are correlated with higher difference values for word duration. Remember that the difference value was obtained by a subtraction of the simple onset word duration from the reduced cluster word duration. A positive value then means that the reduced cluster word was longer than the simple onset word. Adult listeners appear to judge the longer of the two utterances to be the target /Cr/ word. Given that word duration is not a distinctive acoustic characteristic of the /Cr/~/C/ minimal pairs (recall the results from the MANOVA, 2.2.4.1), the participants leaned on the wrong acoustic cue when making up their mind about which of the two utterances in the forced-choice identification task had an onset cluster in its target form.

[10] Word duration refers to the duration of the utterance that was used in the perception test. The utterance could have either a CV or a CVC structure, dependent on whether it was trimmed or not.

The same goes for the second predictor suggested by the model. According to the results, the F2 difference score was also a good predictor of the variation in the participant's judgments. Given that here a negative correlation was found, it means that out of the two utterances the participants perceived, they chose for the one with lower F2 in the vowel onset. If we recall the results from the MANOVA in 2.2.4.2, we can see that this is opposite to the intention of the toddlers, who tend to use the rise of F2 in the vowel onset to "mark" the omitted liquid in their utterance.

2.3.2.2. /kn/~/k/ word pairs

For the /kn/ dataset, the fraction correct perception, averaged over the 37 items (utterance pairs), is 713/1295 = 55.06%, with the worst item scoring 8.6% and the best item scoring 88.6%. For the correlation analysis we are interested in seeing whether a higher correct perception of an item is correlated with one of the acoustic characteristics of nasality or with a measure of a covert contrast like duration, which were measured in advance.

As with the /Cr/~/C/ pairs, we conduct a logistic regression analysis. The response variable is again whether the participant gave a correct or incorrect answer, but the potential fixed factors in the logistic regression are the differences (between cluster and singleton) in mean F1 (dF1), in "vowel complexity" (dComplex), in overall vowel intensity (dIntens), in center of gravity (dCoG), in vowel duration (dDurV), and in word duration (dDurW), and the vowel type, i.e. whether the onset is followed by /i,ɪ/ or by /a,ɑ,o/ (VowTyp)[11].

With only participant as a (random) factor, the intercept of the model becomes +0.20301. This means that the average log odds in favor of a correct choice are

---

[11] In the /i,ɪ/ word pair set, /ɪ/ was the predominant vowel (16 out of 18), while in the /a,ɑ,o/ set the predominant vowels were /ɑ/ and /o/, 8 out of 19 and 10 out of 19, respectively. In this way we could determine our data set as consisting of front vs. back vowels.

+0.20301, i.e. a correct choice is exp(+0.20301) = 1.2251 times more likely than an incorrect choice. This corresponds to the average percentage correct mentioned above (55.06%/44.94%). The 95% confidence interval of the log odds runs from +0.09370 to +0.31273 (i.e. the odds correct could be anywhere between 1.098 and 1.367), which does not include zero, so we have shown that listeners differentiate the average pair better than chance.

Of the six fixed factors, entering dF1 yields the best improvement of the model (χ^2 = 36.985, *df* = 1, p = $1.2 \cdot 10^{-9}$). After adding dF1, adding dComplex gives the best improvement (χ^2 = 22.264, p = $2.4 \cdot 10^{-6}$). After adding dComplex, adding dDurV into the model yields another improvement (χ^2 = 11.844, p = 0.00058), although adding dDurW would have been almost equally good (χ^2 = 9.955, p = 0.0016). After adding dDurV, only VowTyp can be added (χ^2 = 11.970, p = 0.00054).[12]

In the final model, dF1 has a coefficient of -0.00289/Hz (c.i. = -0.00378 .. -0.00202/Hz; p from zero = $1.1 \cdot 10^{-10}$). This means that every decrease of 10 Hz in the mean-F1 difference between cluster word and singleton word increases the odds for a correct score by a factor of exp(0.0289) = 1.0293, i.e. by 2.93% (c.i. = +2.04% .. +3.85%). This means that the lower the F1 of the cluster word is, the more often it is correctly chosen to be the cluster word, and the lower the F1 of the singleton word is, the more often this is incorrectly chosen to be the cluster word. The interpretation is that listeners associate cluster words with a closed vowel realization and singleton words with an open vowel realization.

The second factor in the model, dComplex, has a coefficient of -5.95/s (c.i. = -9.99 .. -1.96/Hz; p from zero = 0.0036). This means that every decrease of 10 ms in the vowel complexity difference between cluster word and singleton word

[12] The cause that dDurV cannot be included in the model, although it had such a good p value when entering it on its own, is that dDurV and dDurW are strongly correlated, so that entering dDurW takes away dDurV's chances.

increases the odds for a correct score by a factor of exp(0.00595) = 1.0613, i.e. by 6.13% (c.i. = +1.98% .. +10.50%). This means that the more complex the vowel is, the more often it is correctly chosen to be the cluster word, and the less complex the vowel is, the more often it is incorrectly chosen to be the cluster word. The interpretation is that listeners associate cluster words with complex vowels and singleton words with simple vowels.

The third factor in the final model, dDurV, has a coefficient of +1.304/s (c.i. = +0.615 .. +2.019/s; p from zero = 0.000266). This means that every increase of 10 ms in the vowel duration difference between cluster word and singleton word increases the odds for a correct score by 1.01% (c.i. = +0.62% .. +2.04%). This means that the longer the vowel in the cluster word is, the more often it is correctly chosen to be the cluster word, and the longer the vowel in the singleton word is, the more often this is incorrectly chosen to be the cluster word. The interpretation is that listeners associate cluster words with longer vowels and singleton words with shorter vowels.

The fourth factor in the final model, VowTyp, has a coefficient of +0.414 for /i,ɪ/ over /a,ɑ,o/ (c.i. = +0.179 .. +0.651; p = 0.00057). This means that /k(n)i,k(n)ɪ/ words have a higher odds for a correct score than /k(n)a,k(n)ɑ,k(n)o/ words by a factor of exp(0.414) = 1.513, i.e. by 51.3% (c.i. = +19.6% .. +91.7%). This means that if a cluster word contains /kni,knɪ/, it is more likely to be correctly chosen to be the cluster word than if a word contains /kna,knɑ,kno/. The interpretation is that listeners differentiate the /kn/~/k/ word pairs containing a front vowel better than they differentiate the /kn/~/k/ word pairs containing a back vowel.

2.4. Discussion

2.4.1. /Cr/~/C/ word pairs

In Study 1 we carried out four acoustic measurements with the /Cr/~/C/ word pairs, namely word duration, vowel duration, and F2 and F3 movement in the vowel onset. It turned out that only the F2 measure showed a significant difference between the word pairs, the F2 slightly rising in the vowel onset of target /Cr/ words. The other three measures did not reliably differentiate between the word pairs.

For adult production it is known that rhotics exhibit a low F3 together with a small F2–F3 separation (Kent & Read, 2002; Lindau, 1985). In adult speech, Gulian (in preparation) found a rise of F2 in the vowel onset only for front vowels following /Cr/ clusters. However, in the child language data the effect of F2 in the vowel onset found here is not restricted to front vowels. We should therefore consider the possibility that the rising F2 in the child language data is not a trace of the target rhotic, but of another sound that substitutes this rhotic. Since F2 is rising, this segment should have a low F2. In English we would find a ready candidate for this substitute, namely the glide /w/. A salient effect of the glide /w/ (in English) on the following vowel is a rising F2 (Cooper, Liberman & Borst, 1951). Further, Chaney (1988) recorded adult-like articulations of target segments /r, l, w, j/ by four-year-olds and found that among these segments /w/ exhibited the lowest F2 value (829–1179 Hz). Chaney reported a large F2 transition from the glide to the vowel for all vowel types but this transition was faster for front vowels. More evidence from the acquisition of English points to the fact that in cluster acquisition, target /tw/ and /kw/ are the first clusters to be realized correctly and that /Cw/ is the designated substitute for word-initial consonant clusters (Smit, 1993, Smith 1981). However, in Dutch the glide /w/ only occurs between vowels, in unstressed syllables, and it is hardly ever used as a substitute for liquids in word-initial clusters – only twice in the CLPF database for example. It is thus not likely that the rising F2 is a trace of a (later) substitute /w/ for the target rhotic segment. Interestingly, the F2 in the children's productions has a frequency comparable to that of the F3 in adult speech. A more likely explanation for the rising F2 in the children's productions, then, is that it acoustically mimics the rising F3 in post-rhotic vowels in adult speech, a strategy that can arise under circumstances of underdeveloped speaker normalization skills. This could also explain why we do not find a rising F3 as a trace of the target rhotic in the children's productions. The rising F2 might thus be a genuine trace of the rhotic in the target /Cr/ cluster, and shows a first discrepancy between the production and perception of adults and children.

The results of the adult perception experiment show another discrepancy, in this case between the cues that the adult participants relied on to make a decision, and the cues that the children provided. When the two words in the forced-choice identification task perceptually differed in duration, the adult listeners tended to choose the longer one as the target /Cr/ word. This choice might be based on an expected compensatory lengthening of the vowel. While this does occur in reduced coda clusters (Gilbers & Den Ouden, 1994), for target onset clusters the children in our study did not systematically use word duration to differentiate between the minimal pairs. This discrepancy between the children's intentions and the adults' interpretation explains the poor performance in the forced-choice identification task.

2.4.2. /kn/~/k/ word pairs

We measured six acoustic cues for the /kn/~/k/ word pairs, namely, mean F1, vowel complexity duration, overall vowel intensity, center of gravity, vowel duration and word duration. We found that target /kn/ words tend to contain a different vowel complexity duration (i.e. a vowel consisting of more than one vowel element) from target /k/ words. Besides, target /kn/ words probably show lower center of gravity than their /k/ counterparts; this is especially true if a low vowel follows. The vowel complexity trace is a child-specific cue to the intended realization of the /n/ that is not found in adult productions (Gulian & Levelt, 2009). A tentative interpretation of the "vowel complexity measure" effect could be that toddlers' first step into learning to master /kn/ clusters occurs through diphthongizing the vowel. However, the exact nature of this trace needs to be explored further.

In the perception test, participants performed only slightly better than chance. The factor that best predicted the participants' correct choice for the /kn/ target word, was the score for difference in mean F1. That is, listeners associate cluster words with a closed vowel realization and singleton words with an open vowel realization. Furthermore, the perception scores appeared to be associated with the vowel complexity score, where listeners associate cluster words with complex vowels and singleton words with simple vowels. The perception scores were correlated with vowel duration too, where listeners associate cluster words with the words containing longer vowels. Finally, listeners appear to differentiate the /kn/~/k/ word pairs containing a front vowel better than they differentiate the /kn/~/k/ word pairs containing a back vowel.

Two of the difference scores (difference mean F1 and difference vowel duration) which the participants thought to point to target /kn/ clusters, did not significantly differentiate between /kn/~/k/ word pairs in the children's productions. As in the /Cr/~/C/ case, then, there might be a discrepancy between the distinctive cues that can be found in the productions of Dutch children and the cues that adults seem to rely on to distinguish between these productions. Again, this discrepancy is a likely explanation for the poor performance of adult listeners in the perception test.

However, in the /kn/~/k/ case there was also one successful predictor of correct perception, namely the difference in vowel complexity, which indeed distinguishes the children's /kn/~/k/ target word productions. In a perception test, listeners correctly associate target /kn/ words with complex vowels and target /k/ words with simple vowels. However, despite the correctly detected complex vowel in cluster words by the adult participants, the overall perception scores were nearly at chance level due to the apparently higher reliance on misleading cues.

The /kn/~/k/ word pairs were divided into a front vowel word set and back vowel word set. The front vowel word set consisted predominantly of word pairs containing /ɪ/, while the back vowel word set consisted of /ɑ/ words on the one hand and of /o/ words on the other hand. The mean F1 of the front vowel word set was 554 Hz, and 737 Hz of the back vowel word set. We found that words like *knikker*, which were reduced to [kɪkəɹ], were easier to associate with cluster words than words like *knakworst* reduced to [tɑkvɔst] or words like *knoopjes* reduced to [koːpəs]. Moreover we found that the listeners were better at associating reduced cluster words with the words containing vowels with low F1 than with words containing vowels with higher F1. This might explain the high scores in the identification test for the /kni/~/ki/ word pairs, despite the fact that our data set does not consist of prototypical front nor of prototypical back vowels.

2.5. Conclusion

In this chapter we analyzed the acoustic characteristics of pairs of toddler utterances that differed in the presence or absence of an onset cluster in the target words they were producing. We looked at minimal pairs of /Cr/~/C/ and /kn/~/k/ target words, which were realized in such a similar way that they received identical transcriptions in the CLPF database. We found that acoustic traces of the omitted consonants were present in the children's productions (Study 1), but that adult listeners were not very good at picking up on these acoustic cues (Study 2). Even though Dutch children's target rhotics in the complex cluster /Cr/ were not perceived by adult transcribers, the children did tend to realize a rising F2 in the vowel onset, which is reminiscent of the rising F3 that we see in adult speech. As for target /kn/ clusters in which the nasal sound was not perceived by adult transcribers, we found that the subsequent vowel in the child's production did show a moving formant pattern, and a lower center of gravity. These new covert contrasts add to those discovered before, for VOT by McLeod et al. (1998) and for omitted target coda consonants by Song and Demuth (2008).

The existence of an acoustic trace sheds light on the locus of the omission in the production model (see Figure 1). Since a phonetic correlate is present, the consonant must have been available in the lexical representation as well as in the phonological surface representation. The consonant must therefore exist at every stage in Figure 1, although its presence in the acoustic representation may be minimal. One could say that the ultimate "deletion" of the consonant takes place in the subsequent perception by the adult listener.

This communication problem is apparently caused by the acoustic "near merger" that the child creates. As mentioned in the introduction, the source for this near merger might lie in an incomplete mapping from phonological surface form to the appropriate set of auditory target cues; for instance, the child might only implement a low F1 cue for the segment /n/ in the cluster, which for an adult observer may constitute nothing more than a slight non-phonemic diphthongization on the vowel. Alternatively, the source might lie in an incomplete sensorimotor mapping from auditory cues to articulatory gestures; for instance, the child may be able to implement a lowering of the velum for the /n/ in a cluster, but not the required tongue-tip gesture. Or it might lie in developmental restrictions on motor behavior; executing all the required articulatory movements for two subsequent consonants within a certain time requires practice, and therefore for some time only a partial realization succeeds. Alternatively, the information about the adult /r/ or /n/ in clusters that the child stored in her underlying form may have been incomplete, e.g. in the form of a phonological feature on the vowel, which at the phonetic level comes with auditory cues that the adult listener cannot interpret. This requires further study.

In the adult perception experiment, the main predictive factor for perceiving a cluster was duration. Relying on duration is a rational strategy because the correct production of a word starting with a consonant cluster is likely to last longer than its minimal pair starting with a singleton consonant. Unfortunately for the participants in the study, the children in this study did not systematically realize the durational aspect of the target clusters. Adult listeners apparently use different cues in order to disambiguate toddlers' productions than the cues that toddlers themselves provide to contrast between words that on the surface sound like minimal pairs. This “mistuning” between toddlers' intention and adults' perception will of course disappear over time because the child's own perception and articulation abilities improve. Being misunderstood by adult listeners might be a trigger to improve on their word productions.

Appendix 1: List of the words used for acoustic analysis in the present study in their target adult form. The words are followed by a phonetic transcription and an English translation.

1. Beertje [beːɹt͡jə] ‘bear’
2. Bomen [boːmə] ‘trees’
3. Boot [boːt] ‘boat’
4. Borstel [bɔɹstəl] ‘brush’
5. Bril [bʀɪl] ‘glasses’
6. Brokje [bʀɔkjə] ‘piece’
7. Bromfiets [bʀɔmfits] ‘motor bike’
8. Brood [bʀoːt] ‘bread’
9. Daar [daːɹ] ‘there’
10. Die [di] ‘that’
11. Dit [dɪt] ‘this’
12. Draait [dʀaːit] ‘it turns’
13. Draaimolen [dʀaːimoːlə] ‘merry-go-round’
14. Draak [dʀaːk] ‘dragon’
15. Drinken [dʀɪŋkə] ‘to drink’
16. Gat [χɑt] ‘hole’
17. Kijk [kɛik] ‘look’
18. Kat [kɑt] ‘cat’
19. Kaak [kaːk] ‘(nonword)’
20. Kinderboerderij [kɪndəɹbuːɹdəʀɛi] ‘children’s farm’
21. Kindje [kɪnt͡jə] ‘little kid’
22. Kikker [kɪkəɹ] ‘frog’
23. Kip [kɪp] ‘chicken’
24. Koom [koːm] ‘(nonword)’
25. Kopen [koːpə] ‘to buy’
26. Knaak [knaːk] ‘old currency (functions as nonword)’
27. Knap [knɑp] ‘clever’

28. Knakworst [knɑkʋɔɹst] ‘frankfurter sausage’
29. Knie [kni] ‘knee’
30. Knippen [knɪpə] ‘to cut’
31. Knikker [knɪkəɹ] ‘marble ball’
32. Knopen [knoːpə] ‘buttons’
33. Knuppel [knʏpəl] ‘base-ball bat’
34. Noot [noːt] ‘nut’
35. Paat [paːt] ‘(nonword)’
36. Praten [pʀaːtə] ‘to talk’
37. Taak [taːk] ‘task (functions as nonword)’
38. Tak [tɑk] ‘branch’
39. Tas [tɑs] ‘bag’
40. Tekenen [teːkənə] ‘to draw’
41. Tik tak [tɪk tɑk] ‘tick tock of a clock’
42. Thijs [tɛis] ‘(proper boy's name)’
43. Tractor [tʀɑktɔɹ] ‘truck’
44. Trein [tʀɛin] ‘train’
45. Trekken [tʀɛkə] ‘to pull’

Appendix 2: List of children producing cluster reduction; (m) stands for boy and (f) stands for girl.

Database children		Children at day-care centers	
/Cr/ cluster reduction	/kn/ cluster reduction	/Cr/ cluster reduction	/kn/ cluster reduction
Cato (f)	**Cato** (f)	Daniel (m)	Daniel (m)
Robin (m)	**Robin** (m)	Herman (m)	Herman (m)
Jarmo (m)		Mick (m)	Mick (m)
Tirza (f)		Taeke (m)	Taeke (m)
		Matteo (m)	Matteo (m)
		Hannah (f)	Hannah (f)
		Lars (m)	Jan (m)
		Gyula (f)	Daniel C (m)
		Pieter (m)	Denim (m)
		Micky (m)	Evio (m)
		Boaz (m)	Amber (f)
			Mauro (m)
			Han (m)
			Julian (m)
			Sterre (f)
			Sem (m)

Appendix 3: List of /Cr/~/C/ word pairs used for acoustic analysis. The pairs consist of a cluster reduction and singleton utterances from children from the database and the day-care center sets. The word pairs from the database set appear in bold.

Utterance number	Cluster word			Singleton word		
Name child	age	Intended production	Actual production	age	Adult target	Actual production
1. **Cato**	1;10	draak [dʀaːk]	[daːk]	1;11	daar [daːɹ]	[daːɹ]
2. **Cato**	2;0	drinkt [dʀɪŋkt]	[dɪːŋkt]	1;11	dit [dɪt]	[tɪt]
3. **Cato**	2;1	drinken [dʀɪŋkə]	[dɪːŋkə]	2;1	dit [dɪt]	[dɪt]
4. **Cato**	2;1	trekken [tʀɛkə]	[tɛkə]	2;0	tekenen [teːkənə]	[teːkənə]
5. **Cato**	2;4	draait [dʀait]	[d:ait]	2;3	daar [daːɹ]	[daːɹ]
6. **Tirza**	1;8	broek [bʀuk]	[buk]	1;8	boek [buk]	[buk]
7. **Tirza**	1;8	broek [bʀuk]	[buk]	1;8	boek [buk]	[buk]
8. **Tirza**	1;8	broek [bʀuk]	[buk]	1;8	boek [buk]	[buk]
9. **Tirza**	1;11	draaimolen [dʀaimoːlə]	[daːi]	1;11	daar [daːɹ]	[daːɹ]
10. **Tirza**	1;11	draaimolen [dʀaːimoːlə]	[daːi]	1;11	daar[daːɹ]	[daːɹ]
11. **Tirza**	2;0	draaimolen [dʀaːimoːlə]	[daːi]	2;1	daar [daːɹ]	[daː]
12. **Tirza**	2;0	draai [dʀaːit]	[daːit]	2;0	daar [daːɹ]	[taː]
13. **Tirza**	2;0	trein [tʀɛin]	[tɛin]	2;0	kijk [kɛik]	[kɛi]

14. **Tirza**	2;2	kraan [kʀaːn]	[kaːn]	2;2	taart [buk]	[taːt]
15. **Tirza**	2;2	trein [tʀɛin]	[tɛin]	2;2	kijk [kɛik]	[tɛi]
16. **Tirza**	2;4	broek [bʀuk]	[buk]	2;4	boek [buk]	[buk]
17. **Tirza**	2;4	broodje [bʀoːtjə]	[boːtjə]	2;4	boot [buk]	[boːt]
18. **Robin**	2;0	drinken [dʀɪŋkə]	[dɪːŋkə]	1;11	die [ti]	[tɪ]
19. **Robin**	2;0	drinken [dʀɪŋkə]	[dɪːŋkə]	1;11	die [ti]	[tɪ]
20. **Jarmo**	2;3	bril [bʀɪl]	[bɪl]	2;2	beertje [beɹt͜jə]	[biːcə]
21. **Jarmo**	2;2	brokje [bʀɔkjə]	[bɔkjə]	2;2	bomen [boːmə]	[boːmə]
22. Boaz	2;0	traktor [tʀɑktɔɹ]	[tɑktɔɹ]	2;0	tak [tɑk]	[tɑk]
23. Daniel	1;6	praten [pʀaːtə]	[paːtə]	1;6	paat [paːt]	[paːt]
24. Daniel	1;6	trein [tʀɛin]	[tɛin]	1;6	Thijs [tɛis]	[tɛis]
25. Daniel	1;6	traktor [tʀɑktɔɹ]	[tɑtɔ]	1;6	tak [tɑk]	[tɑk]
26. Gyula	2;1	brood [bʀoːt]	[boːt]	2;1	boot [boːt]	[boːt]
27. Herman	3;0	brood [bʀoːt]	[boːt]	3;0	noot [noːt]	[poːt]
28. Herman	3;0	praat [pʀaːt]	[paːtə]	3;0	kaak [kaːk]	[kaːk]
29. Mick	2;2	traktor [tʀɑktɔɹ]	[tɑktɔɹ]	2;2	kat [kɑt]	[kɑt]
30. Micky	2;1	traktor [tʀɑktɔɹ]	[tɑktɔɹ]	2;1	kat [kɑt]	[kɑt]
31. Taeke	1;8	traktor [tʀɑktɔɹ]	[tɑktɔɹ]	1;8	kat [kɑt]	[kɑt]

32. Pieter	2;0	trein [tʀɛin]	[tɛin]	2;0	Thijs [tɛis]	[tɛis]
33. Pieter	2;0	brood [bʀoːt]	[boːt]	2;0	boot [boːt]	[boːt]
34. Pieter	2;0	traktor [tʀɑktɔɹ]	[tɑktɔɹ]	2;0	kat [kɑt]	[kɑt]
35. Hannah	2;1	broek [bʀuk]	[buk]	2;1	boek [buk]	[buk]
36. Hannah	2;1	draakje [dʀaːkjə]	[daːkjə]	2;1	taak [taːk]	[taːk]
37. Hannah	2;1	kraan [kʀaːn]	[kaːn]	2;1	taak [taːk]	[taːk]
38. Hannah	2;1	trein [tʀɛin]	[tɛin]	2;2	Thijs [tɛis]	[tɛis]
39. Matteo	2;2	trein [tʀɛin]	[tɛin]	2;2	Thijs [tɛis]	[tɛis]
40. Matteo	2;2	draakje [dʀaːkjə]	[taːkjə]	2;2	taak [taːk]	[taːk]
41. Matteo	2;3	trein [tʀɛin]	[tɛin]	2;3	Thijs [tɛis]	[tɛis]
42. Matteo	2;3	draak [dʀaːk]	[daːk]	2;3	taak [taːk]	[taːk]
43. Matteo	2;3	kraan [kʀaːn]	[kaːn]	2;3	gat [xɑt]	[kaːt]
44. Matteo	2;3	draak [dʀaːk]	[daːk]	2;3	taak [taːk]	[taːk]
45. Lars	2;6	trein [tʀɛin]	[tɛin]	2;6	Thijs [tɛis]	[tɛis]
46. Lars	2;6	kraan [kʀaːn]	[kaːn]	2;6	taak [taːk]	[taːt]
47. Lars	2;6	draak [dʀaːk]	[daːt]	2;6	kaak [taːk]	[taːt]

Appendix 4: List of /kn/~/k/ word pairs used for acoustic analysis. The pairs consist of a cluster reduction and singleton utterances from children from the database and the day-care center sets. The information about the children from the database appears in bold.

Utterance number	Cluster word			Singleton word		
Name child	age	Intended production	Actual production	age	Intended production	Actual production
1. **Cato**	2;4	knie [kni]	[kiː]	2;3	kikker [kɪkəɹ]	[kɪkəɹ]
2. **Cato**	2;0	knie [kni]	[ki]	1;11	kikker [kɪkəɹ]	[kɪkəɹ]
3. **Cato**	2;6	knikker [knɪkəɹ]	[kɪkəɹ]	2;5	kikker [kɪkəɹ]	[kɪkəɹ]
4. **Robin**	2;4	knip [knɪp]	[kɪp]	2;3	kip [kɪp]	[kɪp]
5. **Robin**	2;4	knippen [knɪpə]	[kɪpə]	2;3	kind [kɪnt]	[kɪnt]
6. **Robin**	2;4	knuppel [knʏpəl]	[kɪpɔ]	2;0	kijk [kɛik]	[kɪk]
7. Daniel	1;6	knakworst [knɑkʋɔɹst]	[kɑkvɔst]	1;6	tak [tɑk]	[tɑk]
8. Daniel C	1;6	knakworst [knɑkʋɔɹst]	[tɑkvɔst]	1;6	tak [tɑk]	[tɑk]
9. Daniel C	1;8	knakworst [knɑkʋɔɹst]	[tɑkvɔst]	1;8	tas [tɑs]	[tɑs]
10. Daniel C	1;8	knoopjes [knoːpjəs]	[toːpjə]	1;8	boot [boːt]	[boːt]
11. Evio	2;0	knopen [knoːpə]	[koːpə]	2;0	boot [boːt]	[boːt]
12. Evio	2;0	knip [knɪp]	[kɪp]	2;0	kindje [kɪnt͜jə]	[kɪnt͜jə]
13.Herman	3;0	knaak [knaːk]	[kaːk]	3;0	kaak [kaːk]	[kaːk]
14.Herman	3;0	knap [knɑp]	[kɑːp]	3;0	kat [kɑt]	[kɑt]
15.Herman	3;0	knippen [knɪpə]	[kɪpə]	3;0	kippen [kɪpə]	[kɪpə]
16. Jan	1;10	knakworst [knɑkʋɔɹst]	[kɑkvɔst]	1;10	tas [tɑs]	[tɑs]

17. Jan	1;10	knakworst [knɑkʋɔɹst]	[kɑkvɔst]	1;10	kat [kɑt]	[kɑt]
18. Jan	1;10	knippen [knɪpə]	[kɪpə]	1;10	kippen [kɪpə]	[kɪpə]
19. Denim	2;0	knippen [knɪpə]	[kɪpə]	2;0	kippen [kɪpə]	[kɪpə]
20. Denim	2;0	knopen [kno:pə]	[ko:pə]	2;0	kopen [ko:pə]	[ko:pə]
21. Mick	2;2	knippen [knɪpə]	[kɪpə]	2;2	kippen [kɪpə]	[kɪpə]
22. Mick	2;2	knopen [kno:pə]	[ko:pə]	2;2	kopen [ko:pə]	[ko:pə]
23. Mick	2;2	knakworst [knɑkʋɔɹst]	[kɑkvɔst]	2;2	kat [kɑt]	[kɑt]
24. Taeke	1;8	knippen [knɪpə]	[kɪpə]	1;8	kippen [kɪpə]	[kɪpə]
25. Taeke	1;8	knopen [kno:pə]	[ko:pə]	1;8	kopen [ko:pə]	[ko:pə]
26. Matteo	2;2	knoopjes [kno:pjəs]	[ko:pjəs]	2;2	koon [ko:n]	[ko:n]
27. Matteo	2;2	knoop [kno:p]	[ko:p]	2;2	kopen [ko:pə]	[ko:pə]
28. Matteo	2;2	knoop [kno:p]	[ko:p]	2;2	koom [ko:m]	[ko:m]
29.Hannah	2;1	knoopjes [kno:pjəs]	[ko:pjəs]	2;1	boten [bo:tə]	[po:tə]
30.Hannah	2;1	knoopjes [kno:pjəs]	[ko:pjəs]	2;1	boot [bo:t]	[bo:t]
31.Hannah	2;1	knoopjes [kno:pjəs]	[ko:pjəs]	2;1	toon [to:n]	[to:n]
32. Amber	2;2	knikker [knɪkəɹ]	[kɪkəɹ]	2;1	kikker [kɪkəɹ]	[kɪkəɹ]
33. Mauro	2;0	knikker [knɪkəɹ]	[kɪkəɹ]	2;0	kikker [kɪkəɹ]	[kɪkəɹ]

34. Han	2;0	knikker [knɪkəɹ]	[kɪkəɹ]	2;0	kikker [kɪkəɹ]	[kɪkəɹ]
35. Julian	1;11	knikker [knɪkəɹ]	[kɪkəɹ]	1;11	kikker [kɪkəɹ]	[kɪkəɹ]
36. Sterre	2;0	knikker [knɪkəɹ]	[kɪkəɹ]	2;0	kikker [kɪkəɹ]	[kɪkəɹ]
37. Sem	2;2	knikker [knɪkəɹ]	[kɪkəɹ]	2;2	kikker [kɪkəɹ]	[kɪkəɹ]

Chapter 3. A longitudinal analysis of the production of target words with /Cr/ onset clusters.

3.1. Introduction

In Chapter 2 children's productions of target words with onset clusters were acoustically analyzed, and this led to the discovery of acoustic traces of target cluster segments that appeared to be omitted from the children's productions. In spectograms of target words with an onset plosive+/r/ cluster, like *broek* 'trousers' and *krijt* 'crayon', the second formant, F2, in the vowel following the reduced cluster showed a significant raising compared to the F2 in the vowel of similar target words with singleton onsets like *boek* 'book' and *kijk* 'look'. The conclusion was that the rise of F2 in the vowel onset signified a trace of the rhotic from target plosive+/r/ clusters. The presence of a trace indicated that underlyingly, in the lexical representation, /r/ had been present, and it suggested that this /r/ had also been phonologically encoded during the speech production process. However, at the level of phonetic encoding only the rising F2 characteristic of [r] in the following vowel seemed to have been executed, resulting in an only very partial presence of target /r/ in the surface form. The trace data thus indicated developmental delay with the production of onset clusters at the level of phonetic encoding.

As discussed in Chapter 1, in phonological accounts of cluster development usually two basic developmental stages are posited: an initial stage in which underlying /C_1C_2/ is reduced to a singleton [C] – most commonly to C_1 if the target cluster is /Cr/ - and a second stage in which a complete cluster is realized, either correctly or with substituted segments. The initial stage has been accounted for by positing a CV template (Demuth, 1996), a default setting of a complex onset parameter (Fikkert, 1994), or a high-ranked constraint against complex onsets (Gnanadesikan, 1995; C. Levelt et al., 2000). In these accounts, then, the phonological grammar enforces complete omission of one of the cluster consonants. The acoustic analyses in the previous chapter revealed that a child's (very) partial realization of an actual cluster indeed often leaves the adult listener with the perceptual impression of a cluster reduced to a singleton [C]. This was shown in the perception experiment with adult listeners discussed in Chapter 2, and it is also indicated by phonetic transcriptions of children's renditions of target clusters, even those by trained linguists, where single consonants are transcribed even in cases where studying the spectrogram reveals an acoustic trace of the 'omitted' consonant. This raises the question if an initial stage such as postulated by phonological accounts is justified at all if acoustic data instead of phonetic transcriptions are studied, or if such a stage is a perceptual illusion (Richtsmeier, 2010).

This question formed the starting point for a longitudinal acoustic analysis of the productions of target words with plosive+/r/ onset clusters, henceforth /Cr/, of five children from the CLPF corpus, the result of which is discussed in the present chapter. In Chapter 2, data from four of these children were discussed. Minimal or near-minimal pairs of word productions with perceptually determined singleton onsets were studied, where one member of the pair was the production of a target word starting with an onset cluster and the other member the production of a phonologically similar target word with a singleton onset. In these data, significant F2 raising was found in the vowels following the reduced target /Cr/ clusters compared to those following a similar target singleton /C/. Having established this F2 raising trace on the basis of the restricted minimal pair data, for the longitudinal analysis in this chapter *all* productions targeting an onset /Cr/ onset cluster were studied for the presence of this trace, or any other possible partial rendition or substitution of /r/. This way, I tried to get a more detailed picture of the developmental path of onset cluster production.

In this chapter, first a detailed step-by-step analysis is provided of the development of target /Cr/ onset clusters in the speech of one child, Cato, between the ages of 1;10 and 2;7. It will be shown that, indeed, traceless omissions of target cluster consonants are found in recordings preceding and overlapping with those in which F2 traces are found. The trace then evolves into a more pronounced substitute for /r/, usually [j].[1] Next, the general validity of the developmental stages determined on the basis of Cato's data is illustrated by the longitudinal data of four other children from the CLPF database. Finally, the developmental stages will be discussed in the light of the developing speech production mechanism.

3.2. Data

The data that will be discussed in this chapter come from five children from the CLPF corpus, Cato[2], Robin, Tirza, Enzo and Eva (Levelt 1994; Fikkert 1994) available in Childes/Phonbank (Rose & MacWhinney, 2006). These five children were selected from the corpus because, crucial for the acoustic analysis, the audio files of these children's productions were available in the database. Enzo already had [Cr] productions in his first recording session and his data are included to illustrate the later stages of acquisition rather than the reduction stages, but the other four children all exhibited development in the realization of target /Cr/ cluster words from – transcribed - reduced cluster productions to correct cluster production. I will start out with a detailed analysis of the data of Cato because she produced the highest average number of target /Cr/ words per session of the children showing development from initial [C] to final correct [Cr] productions (see table 1 below).

[1] While for children acquiring English [w] is the most common substitute for target /r/ (McLeod et al., 2001), for children acquiring Dutch it is [j] (Fikkert, 1994).

[2] In CLPF and in the recordings both the full name (Cato) and the diminutive (Catootje) are used.

Table 1: Data on the five children from the CLPF database participating in the analysis.

	Cato	Robin	Tirza	Enzo	Eva
age first session	1;10.10	1;05.11	1;07.09	1;11.08	1;04.12
age last session	2;07.04	2;04.28	2;06.12	2;06.11	1;11.08
number of sessions	16	23	20	16	11
available sessions with sound files	16	22	20	15	10
mean # words first 3 sessions	127	30	30	239	58
mean # words last 3 sessions	479	457	271	539	118
# target /Cr/ words	173	178	120	179	33
# analyzed realizations	169	167	114	178	25

For the analysis both broad and narrow transcriptions of the children's productions of target /Cr/ words were studied. I only studied the realizations of target /Cr/ words if the sound file was of high quality, without intervening voices or noise. In Table 1 the total number of target /Cr/ words and the number of target /Cr/ words that are analyzed are given. The broad transcriptions came from the CLPF corpus in Childes/Phonbank. The narrow transcriptions resulted from an acoustic and perceptual analysis of the sound files of these productions using Praat (Boersma & Weenink, 2008). Each realization of a target /Cr/ onset word was studied visually and auditorily, and special attention was paid to acoustic traces and to acoustic signs that indicated that a different segment than the target was produced. Visually this process was supported by oscillograms and waveforms; auditorily, the selection tools in *Praat* made it possible to zoom in and listen to separate parts of the signal in isolation. In a text grid the transcription from the CLPF database was aligned with the waveform and oscillogram. If the transcription from the CLPF database lacked informative detail, an extra layer was added in the text grid that contained the additional (narrow) information. As could be expected, if discrepancies showed up between broad and narrow transcriptions of a word production, the broad transcriptions showed less complexity than the narrow transcriptions. An example of the result of this procedure is below in Figure 1, where the oscillogram and waveform of Cato's production of the target word *trui* (sweater) is presented, together with a text grid containing both the broad transcription [tɛi] from the CLPF database, on the lower layer, and the added narrow transcription, showing more complexity, on the higher layer.

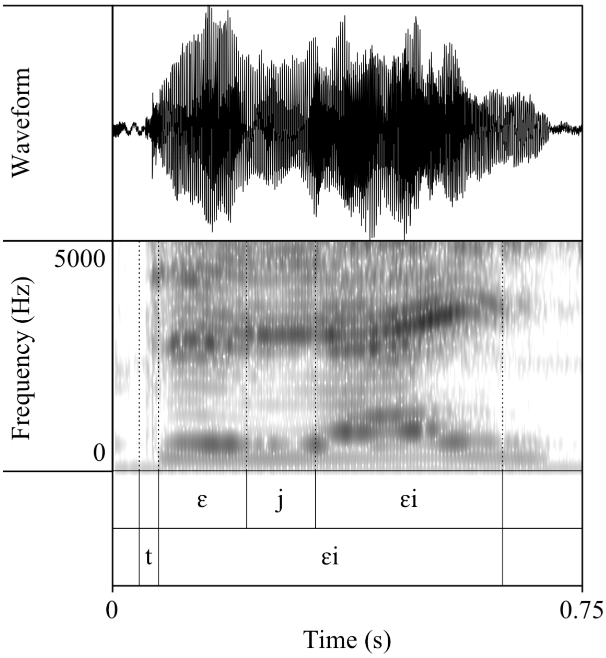


Figure 1. Broad and narrow transcription of Cato's production of the word *trui* (sweater) /trœy/

Based on the narrow transcriptions, for every recording sessions children's cluster realizations were assigned to one of seven different categories:

(1) complete C2 deletion;

(2) C2 deletion, with trace[3];

(3) C2 = glide or vowel;

(4) epenthesis + C2 substitute;

(5) C2 substitute;

(6) epenthesis + (immature) rhotic;

(7) correct.

I will now first illustrate this analysis with data from Cato. The procedure leading to the categorization of Cato's realizations will be shown on the basis of examples from the different categories. The developmental pattern that emerges from this analysis will then be discussed. Subsequently, I will turn to the data of the other children, and compare their development to the developmental pattern shown by Cato. The original transcriptions from the CLPF database and their new narrow transcriptions resulting from the analysis of all the productions of target words starting with a plosive + rhotic from Cato, Eva, Robin, Tirza and Enzo can be found in the Appendixes 1 to 5.

3.3. Cato's development of the production of target /Cr/ onset clusters

Table 2 below shows how often target /Cr/ words occur in Cato's vocabulary in the nine-month recording period between the ages of 1;10.10 and 2;07.04. The percentage of target words starting with /Cr/ does not increase over time. On the contrary, it seems to drop after session 9. The total number of words per session shows a big increase from session 9 on, due to the fact that Cato starts to utter longer phrases and sentences. While before, words with an onset

[3] Except for the F2 raising, the presence of a vocalic transition between the initial consonant and the following vowel was also considered to be a trace of the omitted target /r/.

cluster were often produced in isolation and were repeated several times in the same recording session, from session 9 on they are often embedded in phrases containing function words, which usually do not start with onset clusters. This explains why there are relatively more target /Cr/ words before than after session 9.

Table 2. Total number of words per session and the number and proportion of target onset /Cr/ words in Cato's recordings.

Session number	Age	Total number of words	/Cr/ words (plosive + /r/)
Session 1	1;10.10	125	11 (8.8%)
Session 2	1;10.24	112	9 (8%)
Session 3	1;11.09	146	5 (3.4%)
Session 4	1;11.22	180	9 (5.0%)
Session 5	2;00.06	325	11 (3.4%)
Session 6	2;00.19	158	5 (3.2%)
Session 7	2;01.03	170	10 (5.9%)
Session 8	2;02.14	177	7 (4%)
Session 9	2;02.28	491	17 (3.5%)
Session 10	2;03.25	478	13 (2.7%)
Session 11	2;04.11	440	8 (1.8%)
Session 12	2;04.25	336	11 (3.3%)
Session 13	2;05.08	424	9 (2.1%)
Session 14	2;05.22	506	12 (2.4%)
Session 15	2;06.06	397	18 (4.5%)
Session 16	2;07.04	534	18 (3.4%)

The acoustic analysis that was performed will now be illustrated with a set of spectrograms of different utterances that are representative for both the procedure that was followed and for the different developmental stages that we will encounter. Some of the spectrograms are of minimal pairs that featured in the analysis of the previous chapter. The F2 raising trace that was found on the basis of the minimal pair comparison was present in a statistically significant way in these data, but it was not present in every single production of a target word starting with a /Cr/ cluster. Here these data are studied individually, and the minimal pairs that feature below therefore highlight either the similarity or the difference between the productions of words with target complex and target singleton consonants. In case the vowel-onset patterns of the two productions are very similar, I conclude that target /r/ is completely omitted from the production. This would give support to the existing phonological accounts of onset cluster development, and in terms of speech production would point to a phonological encoding problem - or a lexical specification problem. Alternatively, if the vowel-onset patterns differ, and a residual of target /r/ is found in the production of the vowel-onset of the target word, this would indicate that no complete omission has taken place. The deviant production is then more likely the result of a phonetic encoding problem.

I start out, in Figure 2 below, with a near-minimal pair from Cato's first recording session, namely the initial part *brand* 'fire' /bʀɑnt/ of the target word *brandweerauto*, 'fire-engine', produced as [bɑnt], and the target word *bal* 'ball' /bɑl/, produced as [bɑw].

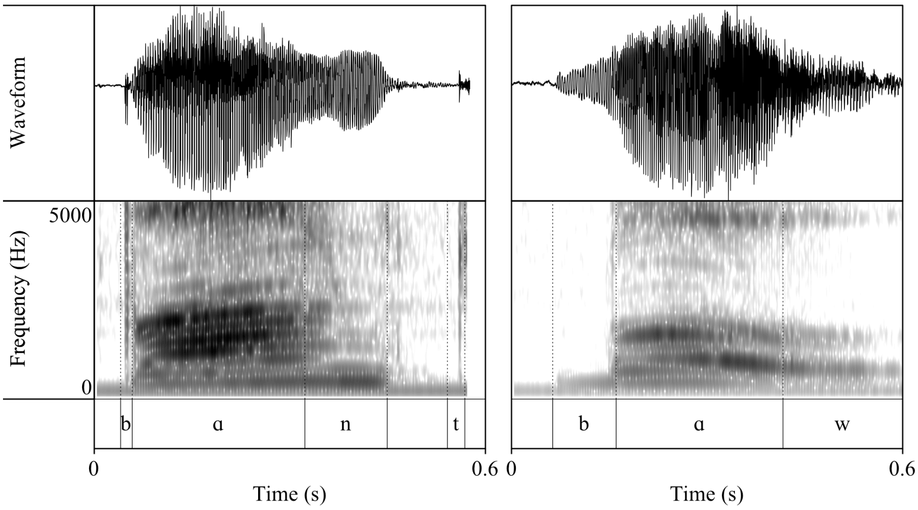


Figure 2: Cato (1;10.10). Left: first part of *brandweerauto (brand)* 'fire-engine' versus Right: *bal* 'ball'

What can be seen in Figure 2 is that although the production of target *brand* has an overall higher intensity than the production of target *bal*, the formant patterns of the vowels are very similar. In the vowel onset the transition from the plosive to the [ɑ] takes a little bit longer in [bɑw] (around 30 ms) than in [bɑnt] (around 20 ms). In both vowels the transition involves a slight rising of the first three formants: in the case of [bɑnt] F1 starts at 725 Hz and reaches 1056 Hz in the vowel nucleus, F2 rises from 1469 Hz to 1552 Hz and F3 1860 Hz to 1960 Hz. Formant transitions are similar in the case of [bɑw]: from 620 Hz to 911 Hz for F1, from 1450 Hz to 1590 Hz for F2, and from 1760 Hz to 1965 Hz for F3. The only difference between [bɑnt] and [bɑw] is that in the latter the F3 is not as stable and clearly visible as in the case of [bɑnt]. Nevertheless, it can be concluded that both formant patterns are similar enough and that there is no evidence for a trace of a different segment in the vowel onset of the reduced cluster word [bɑnt].

In Figure 3 is another near-minimal pair from the same recording session. Here Cato produces [kit] for both the target words *krijt* /kʀɛit/ 'crayon' and *kijk* /kɛik/ 'look'. Since in the first part of both productions of the vowel [i] the F2 is weaker than in the second part, I will first focus on the second part.

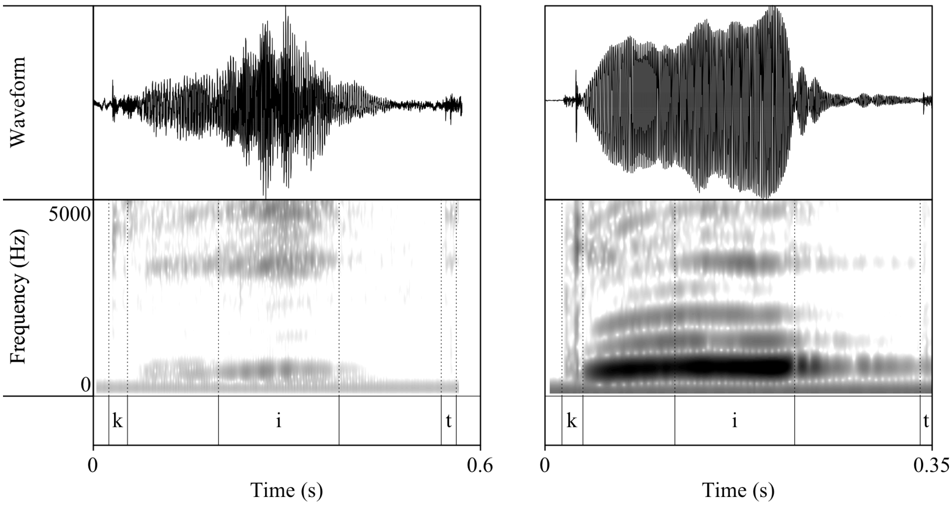


Figure 3: Cato (1;10.10). Left: *krijt/krijtje* 'crayon', Right: *kijk* 'look'

The F1 of target *kijk*, on the right, has a high intensity and is stable around 700 Hz. Above F1 there seem to be two formants that are parallel to F1, but these are in fact harmonics.[4] The formant far above the harmonics, around 3350 Hz, is the actual F2. Immediately above the F2, we can see a weak F3 at around

[4] Auditorily the vowel is clearly [i], and for [i] an F2 of 1400 Hz is too low. A sound with an F1 of 600 Hz and an F2 of 1400 Hz would make an [o] (Weenink, 1985), which is clearly not the case here.

4070 Hz. The [i] of target *krijt* also has F1 around 540 Hz, F2 around 3330 Hz and F3 around 4000 Hz. Here we do not see the harmonics. The first part of the vowel is almost identical to the second but slightly weaker. This is also the case for *kijk*, the only difference being that F2 is not as flat as in [kit] (*krijt*). Again, there are no real differences in the formant patterns in the vowel onset of the two versions of [kit], and I therefore conclude that the vowel onset in *krijt* lacks an acoustic trace of target /r/ and that /r/ thus has been completely omitted.

In Figure 4 is a near-minimal pair from the second recording session, at the age of 1;10.24. Here we have the productions of the target words *draak* 'dragon', on the left, and *daar* 'there', which in the CLPF corpus received the transcriptions [da:k] and [da:]. In the case of *draak*, however, when smaller portions of the vowel were studied, a separate segment could be discerned both visually and auditorily in the vowel onset of *draak*. This is indicated in the narrow transcription on top of the original transcription in Figure 4.

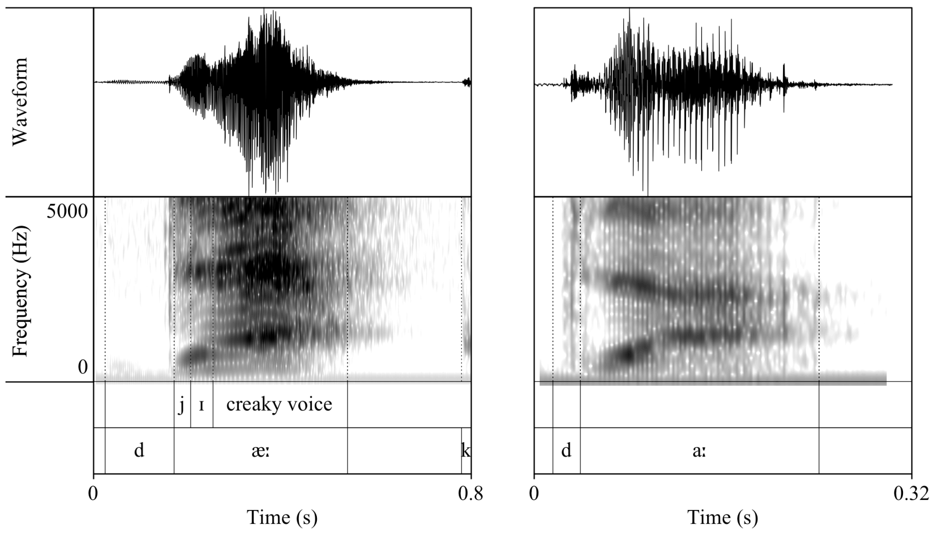


Figure 4: Cato (1;10.24). Left *draak* ‘dragon’, Right *daar* 'there'

In the production of target *daar*, on the right in Figure 4, we see a long transition from the plosive to the vowel [a]; only after 40 ms are the formants of the [a] are expressed. After the vowel onset, F1 rises with a big increase of 400 Hz, while F2 lowers with 400 Hz. F3 cannot be discerned. In the production of target *draak* a transition of around 100 ms takes place before the vowel [a] is uttered. Here F1 and F3 are clearly recognizable while F2 is hard to see. The reason to assume that the formant above F1 is F3, rather than F2, is that its mean value lies around 3050 Hz, too high for F2. Except for the longer transition, a big difference between the acoustic patterns of *draak* and *daar* is that in the first one F2 and F3 are rising towards the vowel nucleus (F2 rises from 2307Hz to 2464Hz and F3 from 2889Hz to 3083Hz), while F3 lowers in the second one. The vowel of target *draak* starts out as [ɪ], and then becomes [æ]. Since this is different from what happens in the vowel of target *daar*, I categorized this utterance as showing a trace of the target /r/.

In Figure 5 on the left is the production of the target word *druif* 'grape' /drœyf/ from the next recording session, at age 1;11.09. A target word with the same diphthong and a singleton coronal onset – *tuin* 'garden' /tœyn/ - from the same recording is on the right for comparison. The transcription of Cato's production of target *druif* in the CLPF database is ['di'jɑuχ̠], already clearly showing that /r/ is being attempted. Looking and listening more carefully, the utterance can be narrowly transcribed as [dɪɛauχ]. The production of *buiten* is transcribed as ['bœi'ti] in the database.

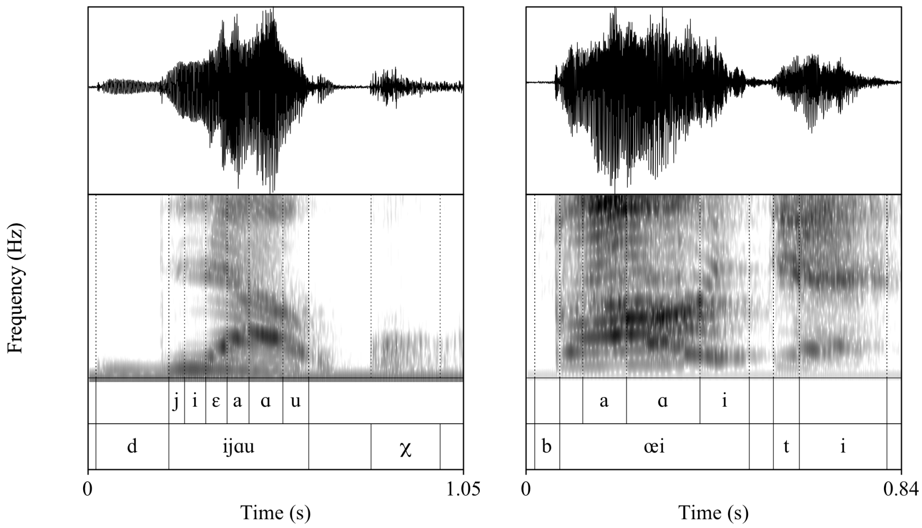


Figure 5: Cato (1;11.09) and (1;10.24). Left *druif* 'grape', Right *buiten* 'outside'

It is clear that Cato does not master the diphthong [œy] yet, and replaces it with [ɑu]. In *tuin*, there is a short transition from the plosive, which lasts around 20 ms, after which high F1 and low F2 are achieved. However, in *druif* the transition from the plosive lasts more than 100 ms. After auditory inspection of this part, it was transcribed as the vowel [ɪ], characterized by low F1, high F2 and an F3 immediately above and hardly discernible from F2. We will see that in Cato's later substitutions for /r/ in target /Cr/ clusters, this [ɪ] becomes a more consonant-like [j].

In Figure 6 are two target words starting with the complex onset /tr/, namely *trui* 'sweater' /trœy/ and *trein* 'train' /trɛin/, from the subsequent recording session at age 2;00.19. The word *trui* 'sweater' is produced twice in this session, once transcribed ['t̠jœ̧ˑi] and once as ['tɪɛˑi] in the database. The word trein is transcribed as [ˌtɪ'jɛin]. Obviously there is still no rhotic present in these utterances, but a substitute vowel or glide can now be clearly perceived.

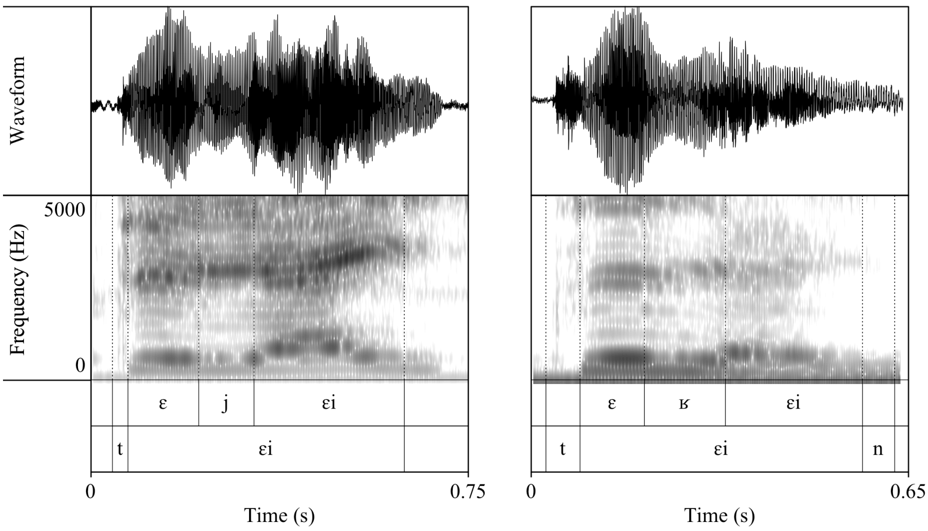


Figure 6: Cato (2;00.19). Left *trui* 'sweater', Right *trein* 'train'

In this particular session, session 6, at the age of 2;00.19, Cato uses [ɛ] + [j] or [ɛ] + [ʁ] to express the rhotic. The use of an epenthetic vowel before the glide is a new strategy of the child. In the word *trein* I have transcribed the element after the schwa as an [ʁ] instead of the transcription [j] in the database, based on the fricative quality of this sound.

From session 10 on, at the age of 2;03.25, Cato's [ɛ] + [j] substitutions of target /r/ have turned into schwa + [ʀ]. The vowel [ɛ] has thus changed into a less salient, central vowel, and the glide, or the slightly more fricative [ʁ], has changed into an immature rhotic. First Cato intends to produce *krijtje* and utters [kəʀɛitʃə] and later, when she intends to say *brandweerauto*, of which we have taken the first syllable *brand* transcribed as [bəʀɑnts] (Figure 7). In the case of krijtje, [ʀ] is visually more consonant-like than in *brand*, but in both cases the sounds have a distinct rhotic quality. The symbol [ʀ] is used to reflect this, even though the way to achieve this quality might have been different from the adult way to realize a rhotic.

As for the use of epenthesis, this is a widely observed phenomenon in the acquisition of clusters (Fikkert, 1994; Greenlee, 1974; Lleo & Prinz, 1996; Smit, 1993). Epenthesis is also one of the main strategies of L2 learners to produce clusters in the L2, when in their L1 clusters are not present (Steele, 2002).
Following Gafos (2002), the epenthetic vowel is probably not a deliberately inserted vowel, but rather the acoustic result of the coordination of the articulations of the two consonants from the cluster. If the articulation of the second consonant is initiated after the release of the burst of the initial plosive consonant, a schwa-like transition between the two articulations results. With practice, the timing of the articulations becomes more and more adult-like, i.e. the articulations will become overlapping, and the schwa-like transition then disappears.

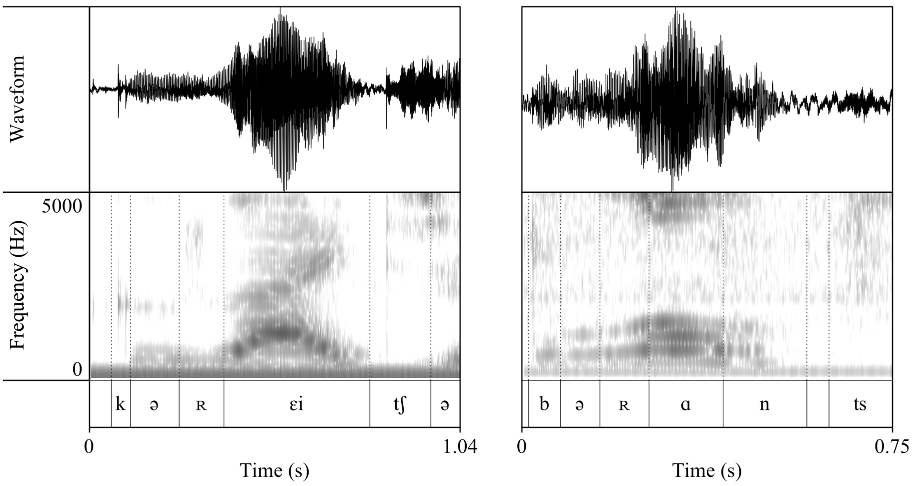


Figure 7: Cato (2;03.25). Left *krijtje* ‘piece of chalk’, Right *brand (weerauto)* ‘fire-engine’

In Figure 8 (left) we see the word *brood* (bread) produced as [bʀoːts], from session 11. A detailed analysis of the utterance shows that the plosive here is not followed by a schwa but rather by a short pause before the rhotic is produced. This illustrates an improved, i.e. shorter transition between the two articulations. In the production of *kreeg* 'received', three months later, on the right in Figure 8 it can be seen that the transition has improved even further, and the release of the burst of [k] now coincides with the onset of [ʀ].

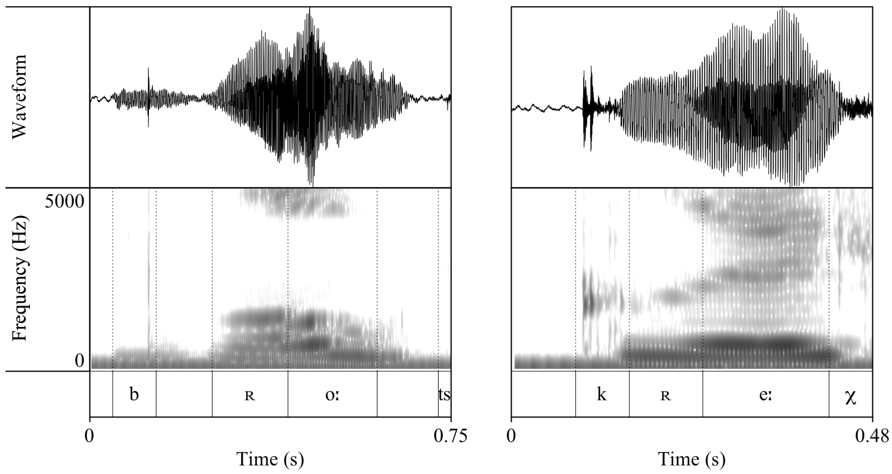


Figure 8: Cato (2;04.11) and (2;07.04). Left *brood* ‘bread’, Right *kreeg* ‘received’

Finally, we focus on the development from an immature rhotic to a more adult-like rhotic with respect to the presence or absence of a trill. To illustrate this, the production of *troon* 'throne' in Figure 9 (right) is used, uttered by the same Cato but now a twenty-three-year old, as a point of reference. The three vertical lines in the spectrogram of the [ʀ] fragment indicate the trill, resulting from the contact between the back of the tongue and the velum.

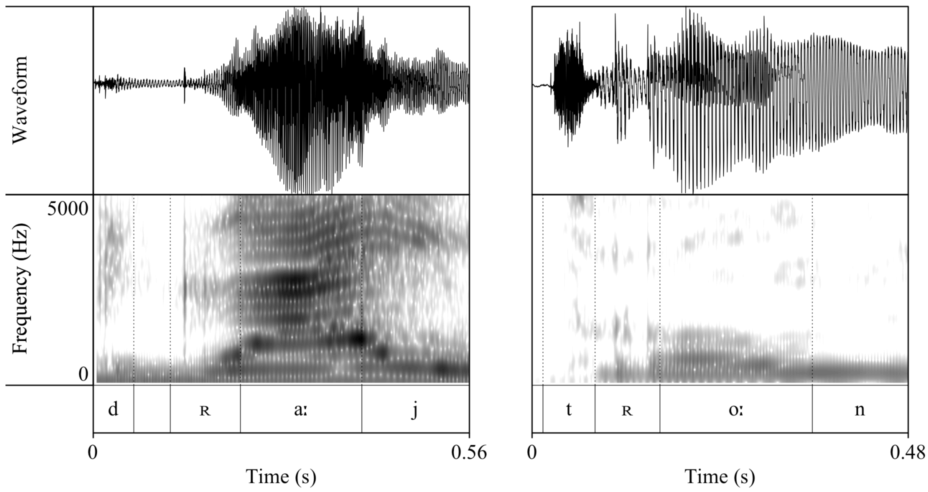


Figure 9: Cato (2;07.04) and (23 years). Left *draai (molen)* ‘merry-go-round’, Right *troon* ‘throne’

In the two [ʀ] productions in Figure 7 and Figure 8 (left) no real trill can be discerned. In the [ʀ] in Figure 8 on the right we can see a single line, indicating a single contact of the back of the tongue with the velum; [ʀ] seems to be a tap produced at the velum. The rhotic in the word draai(molen) 'merry-go-round' produced at 2;07 on the left in Figure 9, finally, is very similar to the adult sound of reference.

In the above section, the qualitative analysis of the target /Cr/ clusters in Cato’s speech has been illustrated. All Cato's productions of target words starting with a plosive +/r/ cluster were analyzed this way, and subsequently scored as instances of one of the seven different categories as discussed in this section and summed up in 3.3.1. The result of this analysis is summarized below in Table 3.

Table 3: Summary of the total number of /Cr/ onset clusters per session and number of their realizations in Cato's recordings.

Session	Age	/Cr/ words	C_2 deletion, no trace	C_2 deletion, trace	C_2 =glide or V	Epenthesis+ C_2 substitution	C_2 substi-tution	Epenthesis + rhotic C_2	C_2 is rhotic
1	1;10.10	11	10	0	1	0	0	0	0
2	1;10.24	9	5	4	0	0	0	0	0
3	1;11.09	5	1	0	2	2	0	0	0
4	1;11.22	9	4	3	1	1	0	0	0
5	2;00.06	9	4	0	3	2	0	0	0
6	2;00.19	5	0	1	0	3	1	0	0
7	2;01.03	10	2	1	0	5	0	2	0
8	2;02.14	6	0	2	2	0	1	0	1
9	2;02.28	16	2	5	3	0	1	3	2
10	2;03.25	12	1	1	0	1	1	3	5
11	2;04.11	8	0	0	0	0	1	1	6
12	2;04.25	11	0	0	1	0	4	0	6
13	2;05.08	9	0	0	0	0	2	1	6
14	2;05.22	12	0	0	0	0	1	0	11
15	2;06.06	18	1	0	0	0	2	0	15
16	2;07.04	18	0	0	0	0	2	0	16

[1] The rationale behind the shadowing is as follows: a single occurrence of a realization is not taken as the start of a stage when it is followed by two or more sessions with zero realizations. A stage ends if a session shows zero realizations of the stage's typical productions for two sessions in a row, and subsequent sessions do not show more than one occasional production.

The shadowing at the upper side of the table helps to show that the different types of realizations make their appearance neatly one after another over time, the correct ones - not surprising - last, from session 8 onwards. The shadowing at the bottom side highlights the disappearance of the particular types of cluster realizations. A new type of realization does not immediately replace the previous type of realization, but rather overlaps with it for some time. In session 9, instances of 6 different realization categories are found. Subsequently, the number of realizations in the first four categories, and somewhat later also those in the fifth category, are quickly reduced to zero. In session 1 we find complete omissions in 10 out of 11 realizations, but in the last three sessions productions are always correct, except for the target word *kroon*, which is always realized as [klon].

On the basis of the analysis presented above, and the summary of our findings in Table 3 above, below in (1) I present the set of partially overlapping developmental stages that Cato passed through on her way to the correct realization of target /Cr/ clusters, focusing on C_2.

3.3.1. Development of /Cr/ onset clusters

Stage 1 (1;10.10-2;02.28): Complete C2 deletion - no acoustic trace
Stage 2 (1;10.24-2;02.28): C2 deletion with an acoustic trace in the vowel
Stage 3 (1;11.09-2;02.28): C2 substitution by a vowel/glide
Stage 4 (1;11.09-2;01.03): C2 substitution + epenthesis
Stage 5 (2;00.19-2;07.04): C2 substitution, no epenthesis
Stage 6 (2;01.03-2;04.11): C2 is (immature) rhotic + epenthesis
Stage 7 (2;02.14-2;07.04): C2 is (immature) rhotic

Target C_2 thus becomes more and more present and then more and more correct in the production of the child over time. Initially it is completely absent from the production (Stage 1); then it shows a minimal presence by an acoustic trace (Stage 2); subsequently it acquires segment status, with a vowel-, or vowel+glide realization (Stage 3, 4); the change to a (correct) consonant realization is accompanied by epenthesis (Stage 6), and finally there is an uninterrupted C_1C_2 sequence (Stage 6, 7). Stage 5, with a C2 substitute realization, is somewhat artificial in Cato's case because it mainly involves realizations of the target word *kroon* "crown", which is steadily produced [klon]. The sequence of productions of target word *krokodil* (crocodile), in (2) illustrates the advancement through different stages over time.

3.3.2. Development of cluster production in *krokodil* 'crocodile'

Stage 1 [kɔkɪdju] – session 5 – 2;00:06
Stage 3 [kuokodɪu] – session 8 – 2;02:14
Stage 4 [kəwoːkədɪu] – session 10 – 2;03:25
Stage 6/7 [kʀoːkodɪl] – session 11 – 2;04:11

Let us now turn to the data of the other children, i.e. Robin, Tirza, Enzo and Eva, and determine to what extent the different realization stages are general developmental stages.

3.4. Developmental stages and the other children

Below I will present a summary of the results of the acoustic analyses I performed on the production data of Robin, Tirza, Enzo and Eva, and some illustrations of these results.

3.4.1. Stage 1: Full deletion

Above it was determined that the production of a target word starting with /CrV/ showed full deletion of C_2 when the spectrogram of the initial CV-part of the utterance was identical to the CV-part of the production of a (semi-)minimal-pair target word starting with /CV/. There is evidence of full deletion of C_2 in the data of Robin, Tirza and Eva.

The recordings of **Enzo** had started at a more advanced developmental period, in which he already produces a (correct) C_2. Throughout the recording period a handful of full deletions of /r/ in target plosive+/r/ sequences is still present though. These deletions involve target words that are also produced correctly, even in the same session, and no clear circumstances that could trigger the deletion can be discerned. They can be seen as relics from stage 1.

In **Robin**'s first 6 recording sessions, up until age 1;07.13, all attempts at target words starting with a plosive+/r/ cluster result in complete omission of C_2. A comparison between Robin's production of the target word *trein* [tɛn] and his production of the word *kijk* [tɛk] at the age of 1;11.02 shows us that the vowel portions in the two utterances are almost identical (see Figure 10). We also see comparable vowels in the CV-parts of the production of *brand* [bɑ] - the first syllable of full *brandweerauto* - at the age of 1;11.20, and the CV-part of *bal* [bɑ] at the age of 1;06.20 (see Figure 11). Robin shows a full-deletion-only period between 1;5.11 and 1;7.27, and after that full deletions keep popping up in his data up until the last recording session, at 2;4.28.

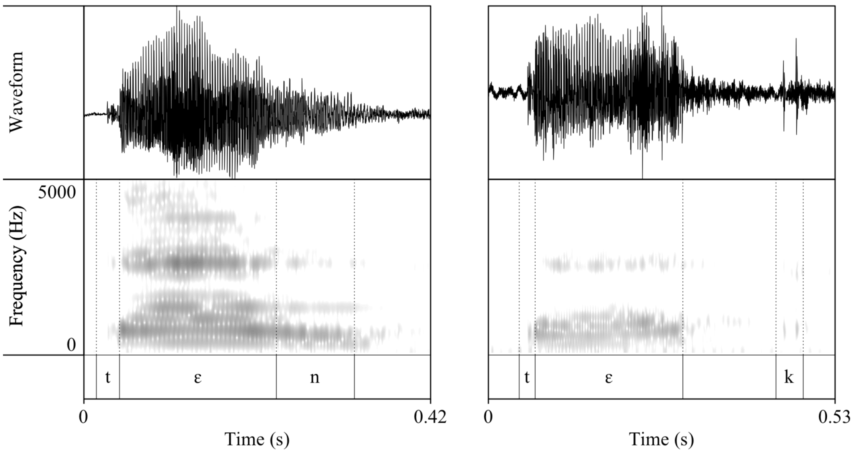


Figure 10: Robin (1;11.12). Left *trein* ‘train’, Right *kijk* ‘look’

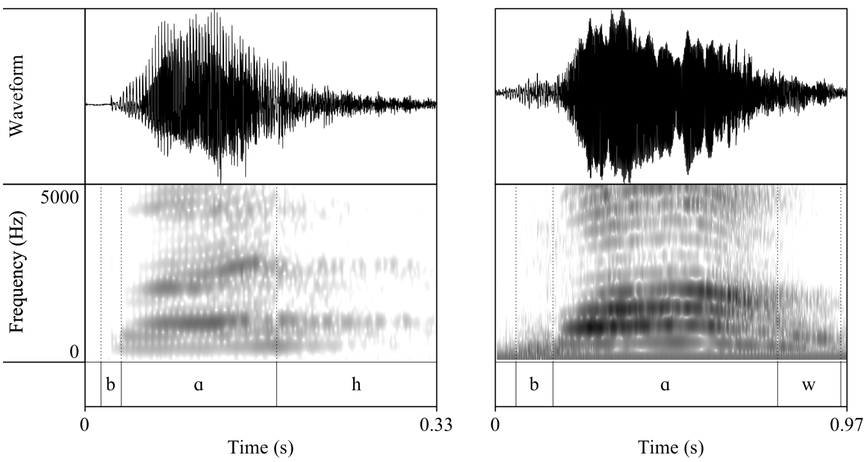


Figure 11: Robin (1;11.20) and (1;06.09). Left *brand(weerauto)* ‘fire(truck)’, Right *bal* ‘ball’

Tirza initially produces very few target words starting with /Cr/. Utterances with full deletions are therefore also only very marginally present in her data.

Between age 1;7.09 and 1;10.08 only four target words are produced, and two of them are produced with a full deletion - of the two others one has an accidentally correct pronunciation, and one a glide substitution of C_2, which otherwise does not occur until much later. The lack of a segmental trace in the acoustic signal is illustrated by the CV-part of the spectrogram of her production of target *kraan* (faucet), produced as [ta:n], which is compared to the initial CV-part of the semi-minimally different *daar* (there), produced as [da:h] at the age of 2;00.05 (Figure 12). The vocalic portions in the two utterances are almost identical, showing that full deletion of C_2 has taken place in target *kraan*. Between 2;00.05 and 2;02.25 full deletion of target /r/ is Tirza's most frequent rendition of a target plosive + /r/ cluster.

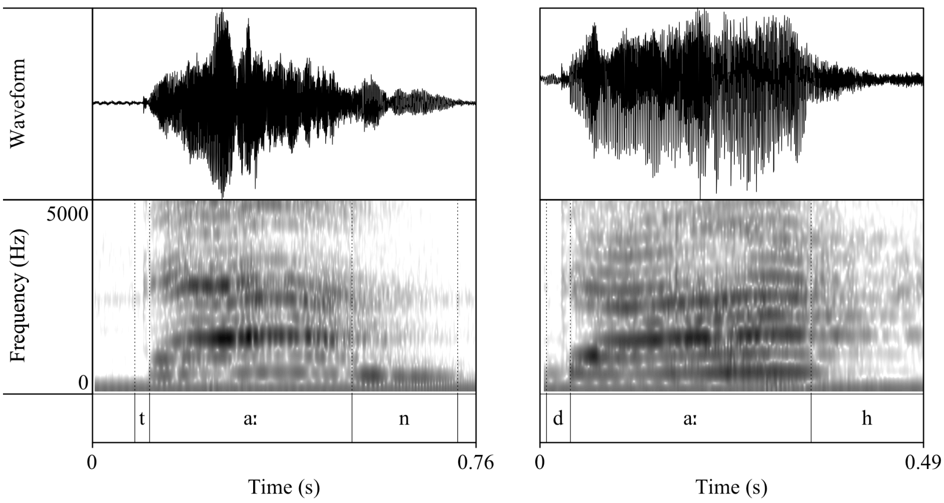


Figure 12: Tirza (2;00.18) and (2;00.05). Left *kraan* 'faucet', Right *daar* 'there'

Finally, **Eva's** productions of target onset /Cr/ clusters also show full deletion of C_2, from the first recording at 1;04.12 up until the pre-final recording at 1;10.03. In her first recordings she produces the word *trein* (train) several times and its vocalic portion resembles the vocalic portion of target *kijk* (look), produced as [tɛit], i.e. there is no trace of /r/. She actually only produces 33 target plosive+/r/ words in her entire recording period, and does not show a lot of development at all.

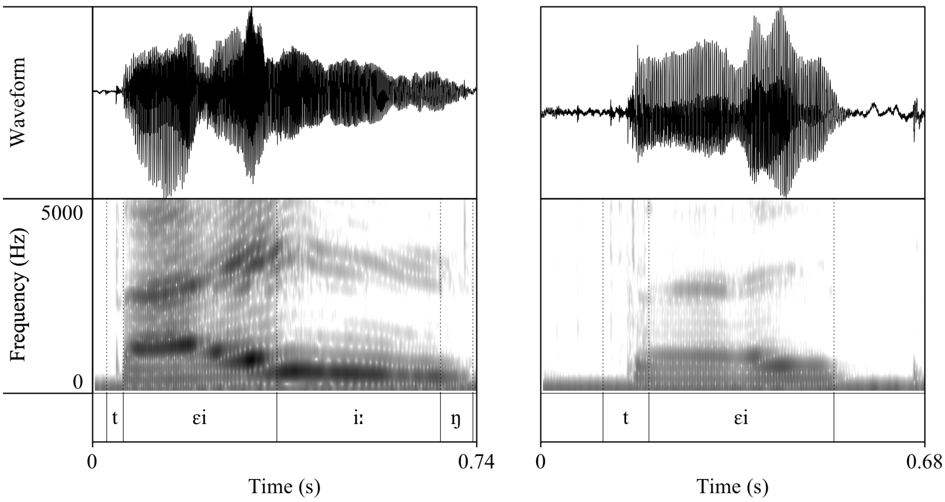


Figure 13: Eva (1;07.15) and (1;06.11). Left *trein* ‘train’, Right *kij (ken)* ‘look’

3.4.2. Stage 2: Deletion with a trace

In Cato’s productions of target /Cr/ clusters, the first sign of the presence of C_2 in the production was a noticeable acoustic trace in the vowel onset of the vowel following C_1. This trace was not present in the 10 productions of target /Cr/ clusters, at age 1;10.10, but appeared two weeks later, at age 1;10.24. Productions with full deletion and productions with a trace then co-occurred in the data for the next four months.

In the data of **Enzo**, being a more advanced cluster producer, no productions with only a trace of C_2 are found.

In the case of **Eva**, productions with a trace of C_2 show up sporadically from session 5 on, at 1;06.11, and up until the end of the recording period, at age 1;11.08. Some productions also show a possible alternative to a trace, namely a sound that could be seen as a merger of the two target consonants. For example, the target word *brood* (bread) /brot/ is produced [mop], and [m] could be seen as the result of merging the labial place of articulation and the non-continuity of target /b/ with the sonority of target /r/. Eva shows no further development in her production of target plosive + /r/ onset clusters until the end of the recording period.

In the data of **Robin**, finally, traces show up from session 8 on, at age 1;08.24. In session 9, at age 1;09.10, seven out of the 14 target cluster words are produced with a trace. After that session some productions with traces are found in every session up until the final one at age 2;03.22. Figure 14 illustrates the presence of a trace in Robin's data.

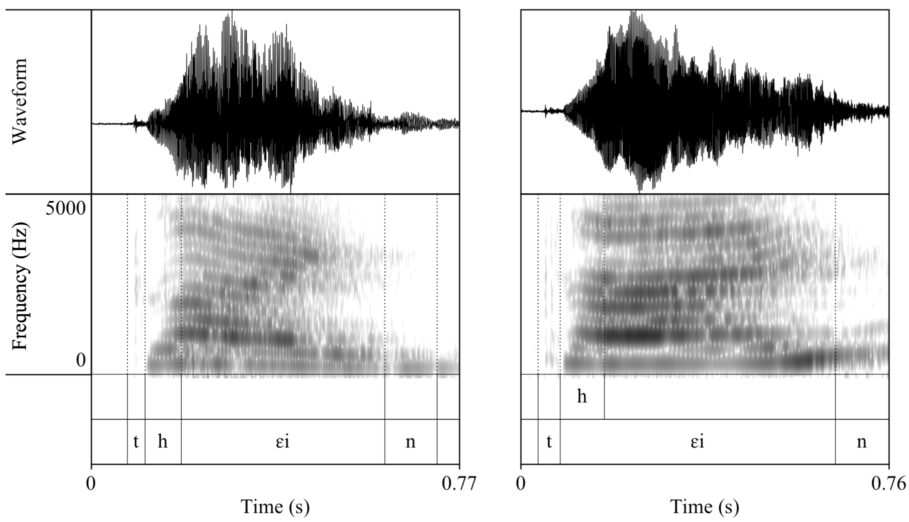


Figure 14: Robin (2;03.09) and (2;02.26). Left *trein* ‘train’, Right *trein* ‘train’

We see that in his realization of the word *trein*, Robin leaves a trace of a vocalic element different from the pattern of the diphthong that follows. This is visible in Figure 14, both left and right. First this is visible in his realization of *trein* at the age of 2;02.26 and later in his realization of *trein* at the age of 2;03.09, where the trace has perceptually taken the shape of noise, transcribed as [h].

In the data of **Tirza** productions with a trace of C_2 are found for the first time in session 6 at 1;10.22. From then on they co-occur with full deletions, and later also with stage 3 renditions, up until session 15, at 2;02;25. In the last four recordings productions with C_2 traces are no longer encountered. In Figure 15 gives an example of a production of target /kro/, with a trace, compared to the production of target /ko/:

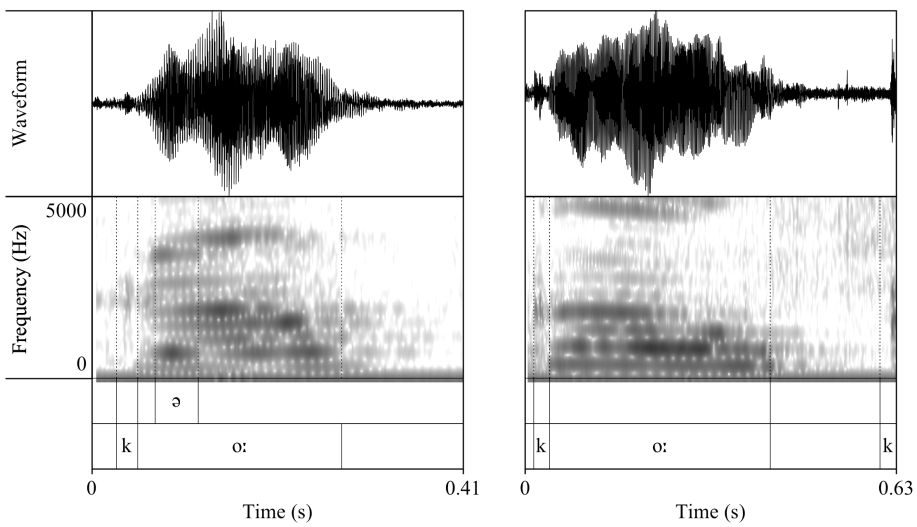


Figure 15: Tirza (2;01.17) and (2;00.05). Left *kro(kodil)* 'crocodile', Right *kook* 'to cook'

As can be seen, a schwa-like element is produced in the vowel onset of target /kro/, which can only be discovered when analyzing the sound file in Praat. This element is interpreted as the acoustic trace of omitted /r/.

3.4.3. Stage 3: C_2 = vowel or glide substitute

In this stage of the development of onset clusters, the child substitutes target C_2 with a glide or vowel. The trace of stage 2 has evolved here into a perceivable, vowel-like segment.

As mentioned above, **Eva**'s development stagnated at stage 2, and she will therefore not be discussed anymore in relation to the subsequent developmental stages. **Enzo**, on the other hand was already beyond this stage by the time his recordings started. The other two children, however, do exhibit examples of stage 3 in their speech.

In the case of **Robin**, stage 3 productions show up for the first time in session 17, at age 2;01.06, almost three months after the onset of stage 2. From then on, stage 1, 2 and 3 productions appear simultaneously up until the last recording session.

In Figure 16 below, we see the realization of the first syllable of the word *brandweer* (*brand*) without an acoustic trace in the signal on the right, and an example of a later realization of this target syllable, with a substitution of the rhotic by schwa on the left. Here the schwa is perceptible even when the word is perceived as a whole.

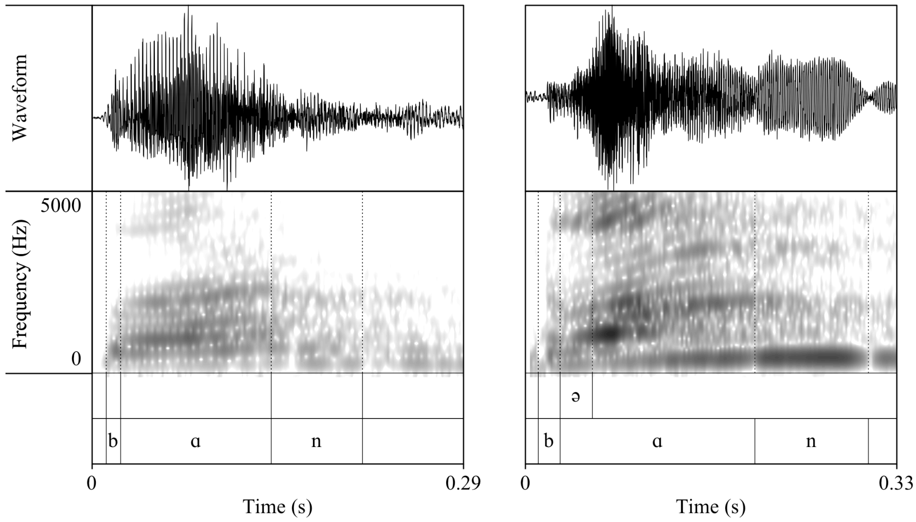


Figure 16: Two realizations of the first syllable of *brandweerauto* ‘fire-engine’, left Robin at 2;01.06, right Robin at 2;02.26.

Tirza produces stage 3 realizations between her recording sessions 11, at age 2;01.02, and 15, at age 2;02.25. They then disappear from her data again. Stage 3 thus forms a stage with a beginning and an end in her cluster acquisition, although stage 1 and stage 2 productions are present in this period too.

In Figure 17 is an example of a substitution of the rhotic by a glide in the first syllable of the target word *trommelaar* 'drummer', on the left, contrasting with the production of the first syllable of a target word with a singleton onset, *potlood* 'pencil' on the right, thus comparing target /trɔ/ with target /pɔ/.

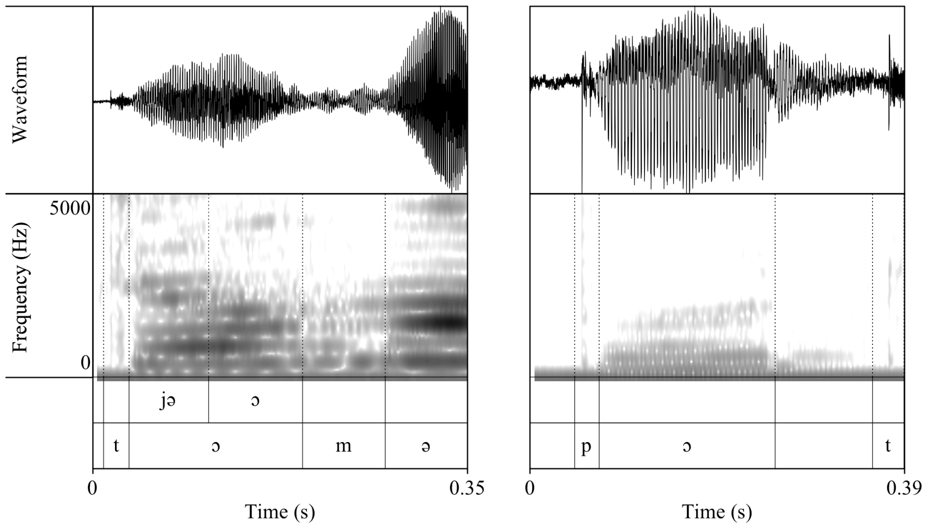


Figure 17: Tirza (2;00.05) and (2;02.12). Left *trom(melaar)* 'drummer', Right *pot(lood)* 'pencil'

3.4.4. Stage 4: Epenthesis + C_2 substitute

In Cato's data productions of target cluster containing epenthesis+C2 substitutions occurred in a subset of the data between the ages of 1;11.09 and 2;01.03. After that period, they completely disappeared from her data. Neither **Tirza** nor **Robin** have any productions reflecting this stage in their development.

For **Enzo**, stage 4 is the first developmental stage we encounter in his data. Stage 4 productions are present predominantly in the first four sessions, between age 1;11.08 and 2;00.13. After that period an occasional realization of this type shows up until the end of the recording period at age 2;06.11. In Figure 18 is Enzo's production of *bramen,* realized as [pəlaumə], illustrating both epenthesis and substitution.

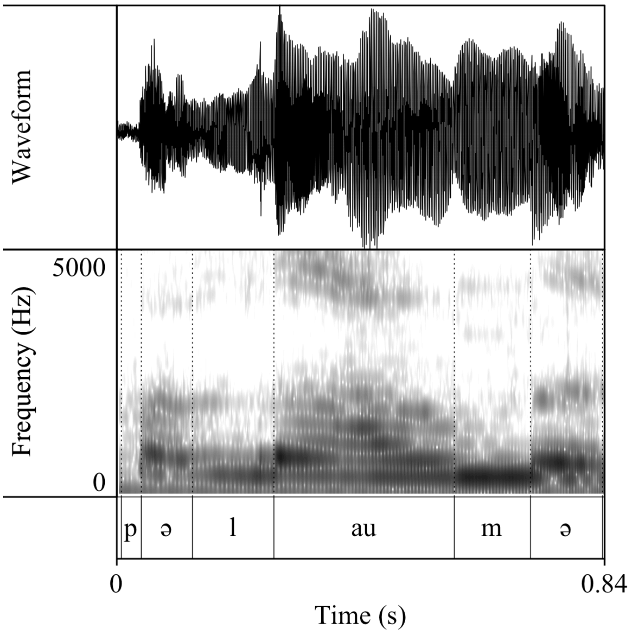


Figure 18: Enzo (1;11.08) *bramen* ‘blackberries’

3.4.5. Stage 5: C_2 substitute, no epenthesis

In Cato's data, the realization of a C_2 substitute without epenthesis applied mostly to a single word, *kroon* "crown", which was produced [klon] up until the last recording session. **Enzo, Robin** and **Tirza** all have a period in which they produce C_2 substitutes, but they all take an individual route to and from this realization phase.

For **Enzo**, this stage collapses with stage 4, epenthesis + C_2 substitution. In other words, Enzo produces C_2 substitutes with or without epenthesis between session 1 at age 1;11.08 and session 10 at age 2;03.14, with a preference for epenthesis in the first 4 sessions, up until age 2;00.13.

Robin directly moves from stage 3, from session 17 at age 2;01.06 on, where C_2 is substituted by a vowel or glide, to this stage 5, where C_2 is substituted by another consonant, in session 20 (2;03.10) and 21 (2;03.22). The two stages co-occur.

Tirza, too, adds stage 5 productions to stage 3 productions from session 13 on, at age 2;02;00. However, in Tirza's case stage 5 really replaces stage 3 from session 16 on, at age 2;03.12 and lasts until the pre-final session at age 2;05.21.

3.4.6. Stage 6: Epenthesis + (immature) rhotic

Cato showed two stages in which the development of a (correct) C_2 realization was accompanied by epenthesis, stage 4 and stage 6. In stage 6, C_2 evolved from a substitute to a(n immature) rhotic. Only Enzo showed stage 4 productions, and only **Tirza** shows productions reflecting stage 6, between session 15, age 2;02.25, and session 18, age 2;05.21. The peak of this stage is at session 16, age 2;03.12, where 7 out of the 9 target plosive + /r/ onsets are realized this way. An example of such a realization is in Figure 19 below, where we see the production [tərɑp] for target word *trap* 'stairs' /trɑp/.

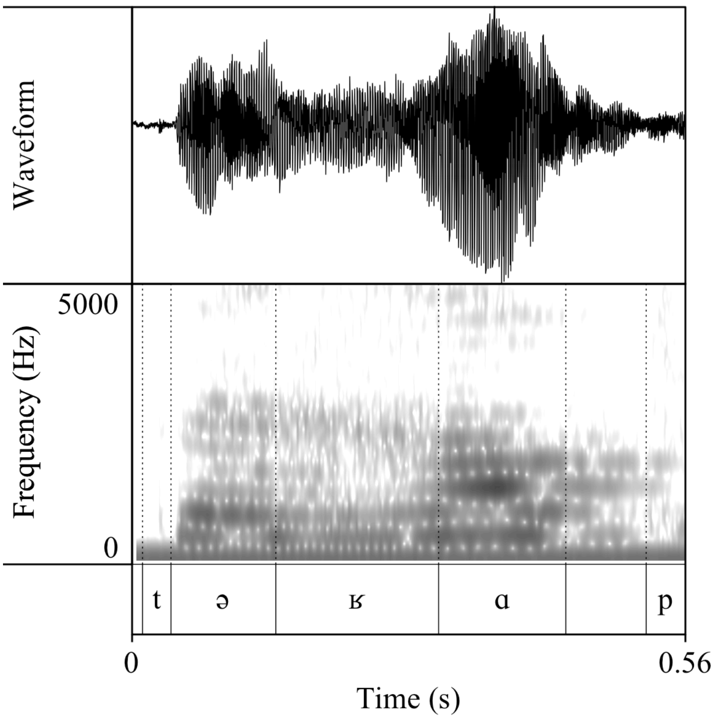


Figure 19: Tirza (2;03.12). *trap* ‘stairs’

3.4.7 Stage 7: C_2 is (immature) rhotic

This final stage in the acquisition of the rhotic in /Cr/ clusters, is attested in the speech of Enzo, Robin and Tirza.

Enzo produces stage 7 realizations from the start of his recording sessions, at age 1;11.08. His rhotics often have a uvular fricative quality ([ʁ]) in the first four sessions of his recordings, but they co-occur with the velar trill realization ([ʀ]). From session 6 on, at age 2;01.03 he exhibits mostly very stable rhotic realizations, which sound adult-like. From session 11 on, at age 2;04.11, hardly any productions from other stages are encountered, so it can be concluded that he has really acquired the plosive + rhotic cluster by that time.

Robin barely reaches stage 7, and only two productions containing a plosive + a rhotic are found, one is session 20, and one in the final session, session 22 at age 2;04.07. The two rhotics he produces in these sessions are both immature, and in the database they are transcribed as voiced uvular fricatives, [ʁ].

Tirza reaches stage 7 in session 17, at age 2;03.27, and only very occasionally produces variants from any other stage in the last two sessions, from age 2;05.05 on. Like Enzo, she has fully acquired plosive + rhotic clusters at this point. In session 17, not all rhotics are adult-like yet, and below in Figure 20, we see an example from Tirza's realization of /Cr/ onset cluster containing such a non-adult-like rhotic. The utterance in Figure 20 is transcribed with the velar trill [ʀ] because auditorily it sounds like a trill. However, it is still immature when compared to the trill produced by Cato as an adult speaker, see Figure 20, right (identical to Figure 9, right). Neither in the spectrogram, nor in the waveform there are clear signs of Tirza's trill, in contrast with the vertical lines, visible in the spectrogram and the vowel-like wave contour in the waveform of the adult production.

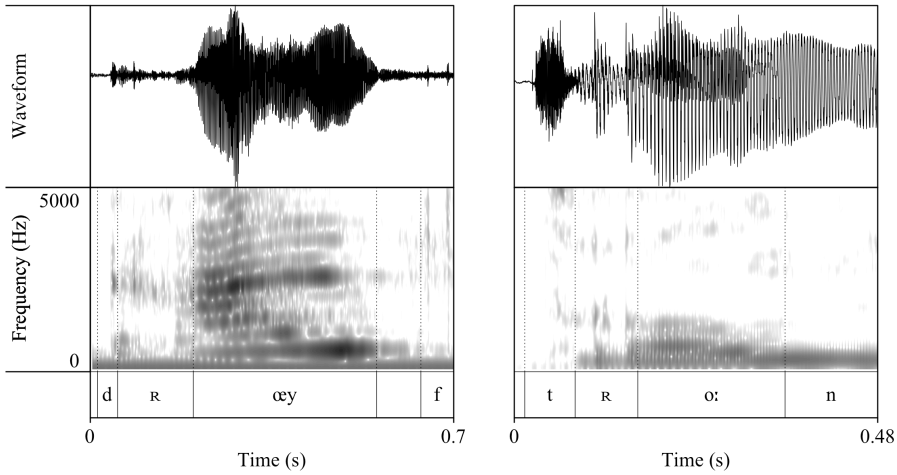


Figure 20: Tirza (2;05.05) and Cato (23 years). Left *druif/druiven* 'grapes', Right *troon* 'throne'

3.5. Summary of the results for all children

In Table 4, below, the developmental steps for all children are summarized. Tables with the individual utterances (broad and narrow transcriptions) for Robin, Tirza and Enzo en Eva can be found in Appendixes 2 to 5.

Table 4. Developmental stages in the realization of target plosive + /r/ onsets

	Stage 1		Stage 2		Stage 3		Stage 4		Stage 5		Stage 6		Stage 7	
Production	Deletion		Deletion + trace		C_2 = glide/V		C_2 = $VC_{substitute}$		C_2=$C_{substitute}$		C_2 =VC_2		C_2=C_2	
start/end session	*S*	*E*	*S*	*E*	*S*	*E*	*S*	*E*	*S*	*E*	*S*	*E*	*S*	*# ses*
Cato	1	10	2	10	3	9	3	7	6	16	7	11	8	*16*
Eva	2	12	5	12										*1s2*
Enzo	1	12					1	10	1	10			1	*16*
Tirza	2	15	6	15	11	15			13	18	15	18	15	*19*
Robin	2	22	8	22	17	22			18	22			20	*22*

In general, the different types of realizations do reflect sequential developmental stages, with a clear starting point, but often a less pronounced end point. That is, the stages are often overlapping, and this will be discussed below in 3.6. Only Cato shows all the seven different types of realizations, appearing successively in session 1, 2, 3, 3, 6, 7 and 8. The stages 1 through 5 all more or less end at the same time, in session 11. Eva and Enzo represent, respectively, the head and the tail of the developmental sequence. Eva only produces utterances reflecting the first two stages of development, while Enzo produces correct onset clusters from the beginning of the recording period on. His development is not so much in terms of successive stages, but in the number of correct clusters he produces, which increases steadily over time. Robin and Tirza show clear successive stages from stage 1 to 7. Robin skips stages 4 and 6, the two stages with epenthesis, and seems to add a new production strategy to the already established ones instead of moving from one stage to the next. In his last three recording sessions, utterances of five different realization types can be found. Tirza also skips stage 4, but does initially use epenthesis with the production of a rhotic. She stops producing stage 1-3 realizations in session 16, and stage 5-6 realizations in session 19.

3.6. Co-occurrence of stages

In the sections above we saw that children could simultaneously exhibit forms from different stages. Robin, Tirza, Cato and Enzo even succeeded to utter forms representing three or more different stages in the same recording session.

We also found instances of full deletions, when complex clusters were produced too. Below are examples of target words produced in different ways within the same session.

Example 1: Enzo, session 6 (2;01.03)

1.1. krokodil *crocodile* /kʀɔkɔdɪl/ [kokodɪl] (Stage 1)

1.2. [kχɔkədiɔl] (Stage 3)

1.3. trein *train* /tʀɛin/ [tʀɛin] (Stage 5)

Example 2: Tirza, session 11 (2;01.02)

draaimolen *Merry-go-round* /dʀaːimoˑlə/

2.1. [tɛiːimoˑlə] (Stage 1)

2.2. [d̩ːɛiˑmoˑlə] (Stage 2)

2.3. [djɛimoˑlə] (Stage 3)

Example 3: Robin, session 20 (2;03.09)

brood *bread* /bʀoːt/

a. [boːt] (Stage 1)

b. [b^hoːt] (Stage 2)

c. [bʁoːt] (Stage 5)

A single word can thus have productions reflecting different stages, or different stages can be reflected in different words. In the data of the five children no clear context can be found that drives the variation in production. For example, in the data it is not the case that words that are used early in the recording sessions are more likely to be produced with a more advanced production than words that have entered the vocabulary at a later stage, nor is it the other way around. It is also not the case that words with a target onset cluster systematically show more advanced stages in single word productions than in multi-word utterances. One possible hypothesis can be formed on the basis of data from Cato, namely that in longer utterances, a cluster - or another more advanced production - is more likely to be produced if the word is situated at the beginning of the sentence than when it occurs at the end of the sentence, as the following examples show:

Example 4: Cato 2;2.15

a. Krokodil ook tanden poetsen   *Crocodile also teeth brush*

['kwoːko'di̠u 'tɑnə 'put̪s̪ə]

   Dit is ook een krokodil   *This is also a crocodile*

[ˌz̠ːə 'oˑkə də 'koːko'di̠u]

b. Krokodil is niet   *Crocodile is not*

['kʋ̣̆oːko'dɯ̆ 'ɩs̪ː 'nit]

   Kijk de krokodil   *Look the crocodile*

['kɛik z̪ə 'k̪oːko'dɯ]

This would point to an effect of attention on the executive functioning of the speech production mechanism. Unfortunately, there are only a handful of such examples available in the data, but it is worthwhile to further test this hypothesis in a controlled production study. Co-occurring or overlapping stages that include traces or substitutions or even correct productions of the target /r/, with or without epenthesis, show that the underlying representation is accurate, and that development is thus taking place at the levels of information encoding and/or the execution of the motor program.

3.7. Discussion

Building on the findings discussed in chapter 2, where acoustic traces of perceptually absent consonants from target onset consonants were found, the first question was if a developmental stage with traceless omissions could be distinguished, preceding a stage with acoustic traces. This was indeed the case. Acoustic analyses showed that in the data of all five children, productions with traceless omissions of target /r/ were present in the data, and in all but Enzo's data these productions occurred before other renditions of target /Cr/ clusters made their appearance. As mentioned in the introduction, in most phonological accounts of cluster development, the initial state of the phonological grammar enforces complete omission of one of the cluster consonants. This initial state can consist of a restricted syllabic template, i.e. a CV template (Demuth, 1996), a default NO setting of a complex onset parameter (Fikkert, 1994), or a high-ranked constraint against complex onsets (Gnanadesikan, 1995; C. Levelt, et al. 2000). In all cases, input word-initial /CC/ combinations will result in [C] outputs of the grammar. The grammatical account could be aligned with processing at the level of phonological encoding. At this level, segments from the lexical representation are grouped into syllables. However, if only a restricted set of syllables can be formed due to a constrained grammar, no cluster spell-out (Levelt, 1989) can take place, i.e. not all segments find a position in the syllable, and they will consequently end up 'stray'. Since only syllabified material is further encoded, stray segments do not show up in production. In addition, because these stray segments are not phonetically encoded, we do not expect to find acoustic traces of these segments in production either. We can thus conclude that an initial stage such as postulated by phonological accounts is justified and that such a stage is not just a perceptual illusion (Richtsmeier, 2010). In the speech production model this account can be incorporated by restricting the syllable spell-out to single onset consonants, leaving no room for /r/.

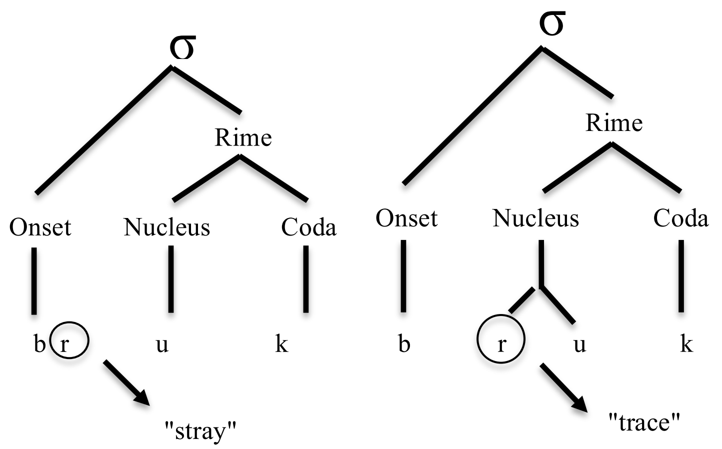
An alternative account is that the omission already occurred at the level of the lexical representation. In a perception experiment where 14-month-olds were exposed to new words containing a coda consonant, like /pat/, it turned out that while participants did perceive this coda consonant, they did not seem to care if they were confronted with a production in which this consonant was omitted (Levelt, 2012). Eighteen-month-olds did notice the omission, and it was concluded that 14-month-olds do not build detailed representations of coda consonants yet. The same could apply to the lexical representations of target /Cr/ onset clusters, where target /r/ might not have been stored in detail (or at all). Since the data analyzed in the present chapter were recorded between 1989 and 1991, it was impossible to check this hypothesis with the five children in this study. Chapter 5 discusses a cluster perception experiment with two-year-olds where this hypothesis is explored further. However, due to fact that results of that experiment were mixed, and the fact that the children in the experiment were older than the children analyzed here in their initial production stages, it is as yet not possible to choose between the two possible accounts of traceless omissions.

One of the main findings in this chapter is that four out of the five children go through a stage where their reduced productions - i.e. productions where target /Cr/ is reduced to [C] - exhibit an acoustic trace of the omitted segment. In the current chapter the term *trace* was used to refer to an acoustic event in the vowel onset which is not normally present in that vowel, but which is not perceivable as a separate segment in the word either. Only after a detailed inspection of the word's oscillogram and spectrogram, plus the use of a narrow phonetic transcription, this trace could be revealed. This stage was found to follow the stage of full deletion of target /r/ and to precede the stage of genuine substitution of this /r/. This stage has not been described before, although other so-called covert contrasts in child language have been noticed (Scobbie, 1998). The trace signifies that the lexical representation must contain information that differentiates the target adult CrV sequence from a target adult CV sequence, that is identical except for /r/.

How can this stage be accommodated in a developing speech production model? The syllable spell-out is still similar to before: there is no cluster spell-out, and only singleton onsets can be formed. However, /r/ is no longer simply left out completely, but it is spelled-out in the nucleus position, together with the following vocalic segment. During phonetic encoding it is accommodated to this vocalic position, resulting in a very partial presence. Whether traces are present or not could be due to factors of planning and attention, i.e. depending on how many attentional resources are available to spell out this 'intruding' segment /r/ in the nucleus position. The two initial stages of phonological encoding are illustrated in Figure 21 below with the target word *broek* 'trousers' /bruk/.

	Stage 1	Stage 2
Lexical representation:	/bruk/	/bruk/

Phonological encoding Syllable spell-out:



Phonetic encoding:

Figure 21: Phonological and phonetic encoding in Stage 1 (traceless omission) and Stage 2 (omission with trace)

In Stage 3, the target /r/ segment can be more fully accommodated in the nucleus position, and this results in the production of a full vowel, like in Cato's [dɪɛauχ] for *druif* 'grape' at age 1;11.09. The nucleus, rather than the onset thus becomes complex first.

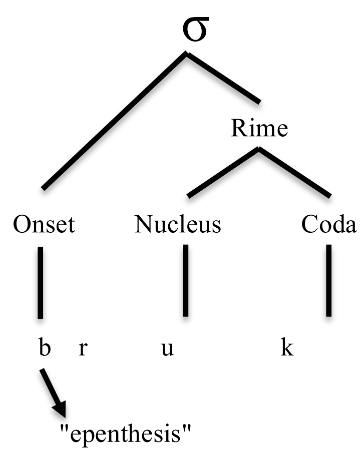
In stage 4 target /r/ is realized as a consonant. However, vowel-epenthesis occurs in-between C_1 and C_2. In phonological accounts, this epenthesis is taken to indicate that the grammar still does not allow for complex onsets. Instead of deleting the C_2, however, insertion of a schwa after C_1 creates a well-formed sequence that includes C_2: C_1VC_2V. In OT, deletion is a violation of the faithfulness constraint MAX, while insertion is a violation of the faithfulness constraint DEP. The development from C_1V in stage 1, to C_1VC_2V in our stage 4, in this framework, then, is due to a re-ranking of these faithfulness constraints, while the Markedness constraint *COMPLEX-ONSET is still in its initial high-ranked position (Gnanadesikan, 1995). However, as already mentioned in section 3.3, it is more likely that the developmental schwa in between the two consonants of a cluster is the acoustic result of an immature coordination of their articulations, namely when the articulation of the second consonant is initiated after the release of the burst of the initial plosive consonant (Gafos, 2002). This would entail that at the level of phonological encoding, a complex onset can now be formed. At a lower level, whether in the encoding of the articulatory plan, or in the actual execution of the plan, immature coordination of the consonant articulations results in a perceived epenthetic vowel.

Stage 4

Lexical representation: /bruk/

Phonological encoding

Syllable spell-out:



Phonetic encoding: [b__ruk]

Figure 22. Phonological and phonetic encoding, stage 4

From stage 4 on, then, the source of deviating productions of target /Cr/ is no longer at the level of phonological encoding. If present in the segment inventory of the native language, rhotics are universally late acquisitions by children acquiring that language (McLeod et al., 2001), and are often substituted by other consonants for prolonged periods of time. In their substitutes, our Dutch children would in general aim for either the dorsality and friction of the Dutch uvular /ʀ/, resulting in [χ], or for a liquid without the trill, resulting in [l]. While this deserves further attention, these approximations of an articulatory goal are likely to result at the level of phonetic encoding. Coordination of the articulations of C_1 and substituted C_2 can improve over time with practice, and epenthesis will disappear (stage 5), but coordination of C_1 with the correct C_2 rhotic can be off again for some time (stage 6).

3.8. Conclusion

In this chapter I have presented a detailed analysis of the acquisition of /Cr/ clusters, by means of descriptive and qualitative analyses of longitudinal data of five children. Seven developmental stages were discovered, but not all children pass through all seven stages, and stages can be overlapping. All children start with the complete and traceless omission of target C_2. This omission occurs during the phonological encoding, where the syllable spell-out only recognizes singleton onsets. Traces and subsequent vowel substitutes of target /r/ result from spell-out of this segment in the syllable nucleus. Spell-out of /r/ in the C_2 position of the onset is the next development. Epenthesis is the acoustic effect of immature coordination of the two consonant articulations at the level of phonetic encoding or during execution of the motor program.

The naturalistic and longitudinal data studied in this chapter provided a clear developmental pattern, but have also left us with questions that can only be answered by performing controlled experiments. In the next two chapters I will turn to experimental methods. In Chapter 4 I will try to confirm - and expand on - the findings from this chapter concerning the development of the speech production model in a production experiment, while in Chapter 5 I report on an perception experiment aimed at determining the amount of detail present in the lexical representation of target words with onset clusters. Detailed studies are still needed to address the question why stages can be overlapping, and to determine the factors that lead to more, or less advanced productions.

Appendix 1: Broad and narrow transcriptions* of Cato's productions of words with /Cr/ onset clusters.

age	session	orthography	broad transcription	narrow transcription
1;10.11	1	trein	te·n	tein
1;10.11	1	trein	theːĭn	tjein
1;10.11	1	trui	tœɪ	tɛːi
1;10.11	1	trui	tɛi	teːi
1;10.11	1	broek	buk	buk
1;10.11	1	broek	buk	buk
1;10.11	1	krijtje	kɑːti	kɑeti
1;10.11	1	krijtje	keːti	keːti
1;10.11	1	brandweerauto	bɑːntɔtʌ	bɑːntɔtʌ
1;10.11	1	brandweerauto	bantʌto·to	bantʌto·to
1;10.11	1	brandweerauto	bantəoto	bantəoto
1;10.24	2	bril	bɛi	bʲɛi
1;10.24	2	draak	da·k	dʲa·k
1;10.24	2	draak	dæːk	dak
1;10.24	2	trein	tɛiːn	tʲɛiːn
1;10.24	2	krijtje	kɛːtɪ	kʲɛit
1;10.24	2	broek	bukʌ̈	bukʌ̈
1;10.24	2	brandweerauto	bɑːç̠əoːtoː	bantəoːto
1;10.24	2	brandweerauto	bãːtəoːtoː	bantəoːto
1;10.24	2	brandweerauto	bɑ̠ːntɪo·to·	bantəoːto
1;11.09	3	bril	bɪjɛi	bɪjɛː
1;11.09	3	druif	dijɑuχ	dʲɛɑuχ
1;11.09	3	krant	kiɑ̠nt	kiɑnt

1;11.09	3	trein	tje˔ĭn	tjɛin
1;11.09	3	trein	tɪj̆ɛin	tɪjɛin
1;11.22	4	bril	ʔmbijɛ˔i	bɪjɛi
1;11.22	4	brood	b˔oːt	bʷoːt
1;11.22	4	brood	ʔmboːt	bʷoːt
1;11.22	4	broek	buk	buk
1;11.22	4	broek	buk	buk
1;11.22	4	broek	buk	buk
1;11.22	4	trein	tɪjɛin	tʷɛin
1;11.22	4	trui	tjœ˔i	tjœi
1;11.22	4	trui	tiɛ˔i	təjɛi
2;00.06	5	brandweerauto	pɑ̪ntəoto	pɑntəoto
2;00.06	5	broemfiets	bɔnɦitʰ	bɔnhɪt
2;00.06	5	draak	djak	dʒjaːk
2;00.06	5	krokodil	kokɪdju˔	kɔkɪdiu
2;00.06	5	krokodil	kɔkokɪɔ	kɔkokiɔ
2;00.06	5	krijtje	kĭjɛitjə	kɪjɛi
2;00.06	5	krijtje	kjɛitə	kijɛi
2;00.06	5	trein	tijɛin	tʒjɛin
2;00.06	5	trein	tjːɛin	tjɛin
2;00.19	6	bromfiets	bɔmə	bʷɔmə
2;00.19	6	trui	tejɛi	təejɛi
2;00.19	6	trui	tĕjɛi	təjɛi
2;00.19	6	trein	tɪjɛin	təjɛin

2;00.19	6	trein	tɪjɛin	tə^{w}ɛin
2;01.03	7	broek	buk	p^{h}uk
2;01.03	7	brandweerauto	bjɑntəoto	bəjɑntəoto
2;01.03	7	brandweerauto	zəbjɑntəoˑto	bɪjɑntəoto
2;01.03	7	bril	bɪːu	bɪu
2;01.03	7	drinkt	dɪŋk	dɪŋk
2;01.03	7	kraai	zəkjaˑŭjă	kəjaːja
2;01.03	7	trein	teːjɛin	tejɛin
2;01.03	7	trein	tɪjɛin	dʒəʁɛin
2;01.03	7	trein	tɪjɛin	təʁɛin
2;01.03	7	trui	tɪjœy	tə^{w}œy
2;02.14	8	brandweerauto	bɑːntəʔoŭto	bəɑntəoˑto
2;02.14	8	drinken	dɪŋkə	dəɪŋkə
2;02.14	8	krokodil	kuoˑkodiu	kuoːkoːdiu
2;02.14	8	krokodil	kʁ̆oːkodɪu	kʁoːkoːdiu
2;02.14	8	krokodil	koːkodiu	k^{w}oːkodiu
2;02.14	8	krokodil	kwoːkodiu	kwoːkodiu
2;02.28	9	brood	bʀoˑt	buʀoˑt
2;02.28	9	bril	bɪuɪ	bɛiw
2;02.28	9	broek	buk	bwuk
2;02.28	9	broek	buk	buk
2;02.28	9	broemfiets	bə̰̃nthitsː	bəɔ̃nthits
2;02.28	9	kraai	kaˑi	kaːi
2;02.28	9	weggekropen	kloːmə	kəloːmə
2;02.28	9	krokodil	kʁ̆oːkodɪŭ	kʁoːkodɪu

2;02.28	9	krokodil	ko:kodıu	kʁo:kodıu
2;02.28	9	kroontje	ko:ntjə	kʁo:ntjə
2;02.28	9	kroontje	klo·ntjə	klontsjə
2;02.28	9	krijtje	kʉɛitjə	kəɛitsjə
2;02.28	9	aantrekken	ʔɑ·nt·ɑĕkə	ɑ:tʁɑ·kə
2;02.28	9	trap	t:ʁ:ɔp	tuɑʁɔp
2;02.28	9	tui	tʁœ·i	tʁa:i
2;02.28	9	draait	d:ai	dəa:i
2;02.28	9	draait	daĭ	dəa:i
2;03.25	10	broek	bř̌u	bʁu
2;03.25	10	brand	bʁɑnt	bəʀɑnt
2;03.25	10	brand	bʀɑnt	bʀɑnt
2;03.25	10	trekt	tʀæ·t	tʀəæ·t
2;03.25	10	trap	tʁɑp	təʁɑp
2;03.25	10	trap	tʁɑp	tʀɑp
2;03.25	10	trein	tʀɛin	tʀɛin
2;03.25	10	gekregen	zəke:ɣə	keje:ɣə
2;03.25	10	kreeg	keχ̌t	çe:xt
2;03.25	10	krijtje	kəʀɛi:tsös	kəʀɛitsɔs
2;03.25	10	krijtje	kʁ:ɛitsə	kəʀɛitsə
2;03.25	10	krokodil	kʁokədıu	kəwokə
2;04.11	11	brood	bʀo:t	bʀo:tsə
2;04.11	11	bril	bʀʉɯ	bʀʉl
2;04.11	11	draaien	draĭjə	dʀa:i

2;04.11	11	kroonje	klo˙ntsə	klontsə
2;04.11	11	krokodil	kʀoːkodɪl	kʁoːkodɪl
2;04.11	11	trap	tʁɑp	təʀɑp
2;04.11	11	trappetje	tʀɑpɑ̯tçə	tʀːɑpətjə
2;04.11	11	aantrekken	ʔantʁɛkə	antʁɑkə
2;04.25	12	broek	ʋuk	bwuk
2;04.25	12	bril	bʁʉɔχ	bʁʉw
2;04.25	12	bruno	bʀʏno˙	bwuno
2;04.25	12	bruno	bʀʏno˙	bʁʏno
2;04.25	12	bruno	bʀʏno˙	bʀʏno
2;04.25	12	bruno	bʀʏno	blɪno
2;04.25	12	bruno	bʀʏno	bʀʏno
2;04.25	12	drie	dʁi	tʁi
2;04.25	12	krullen	kʏwə	gjuwə
2;04.25	12	krullen	kʁʏwə	kʀʏwə
2;04.25	12	kroontje	kloːntjə	kloːntjə
2;05.08	13	bromtol	bʀɔmtɔ	bʀɔmtɔ
2;05.08	13	bril	bʀɪɔ	bʀʉw
2;05.08	13	draak	dʁaːk	dʀaːk
2;05.08	13	drinken	dʀɪŋkə	dlɪŋkə
2;05.08	13	traan	dʀan	dʀamə
2;05.08	13	traantjes	dʀantjəs	dəʀaːtjəs
2;05.08	13	trap	tʁɔp	tʁɔp
2;05.08	13	praten	pʀaːtʉ	pʀaːtə
2;05.08	13	kroontje	klo˙ntjə	tlontjə

2;05.22	14	broek	bʀuk	bʀuk
2;05.22	14	bromtol	bʀɔmtɔ	bʀɔmtɔ
2;05.22	14	afgebroken	ɑfmbʀoːkə	ɔfχəmbʁoːkə
2;05.22	14	bracht	bʀɑχt	bʀɑnt
2;05.22	14	drukken	dʁʏkə	dʁʏkə
2;05.22	14	droog	dʀoːx	dʁoːx
2;05.22	14	gekregen	ŋkʀeːχə	əkʀeːχə
2;05.22	14	prullebak	plʉləbɑk	plʉləbɑk
2;05.22	14	prins	pʀɪns	pʀɪns
2;05.22	14	uitgetrokken	œʏtʀɔkə	œʏtʀɔkə
2;05.22	14	trui	tʀœi	tʀɛi
2;05.22	14	trommel	tʀɔmoŭ	tʀomow
2;06.06	15	gebroken	mbʀoːkə	əmbʀoːkə
2;06.06	15	bromfiets	bʀɔmhits	bʁɔmfits
2;06.06	15	brandweerauto	brɔntʋiɹoːto	brɔntʋiːoːto
2;06.06	15	druiven	dʀœʏfə	dʁœʏfə
2;06.06	15	drinken	dʀɪŋkə	tʀɪŋkə
2;06.06	15	prullebak	plʉləbɑk	plələbɑk
2;06.06	15	praat	mpʀaːt	pʀaːːt
2;06.06	15	precies	pɑsis	pɑsis
2;06.06	15	krant	kʀɑnt	kʁɑnt
2;06.06	15	krant	kʀɒnt	kʀɔnt
2;06.06	15	krukje	kʁʏkjə	kʁʏkjə
2;06.06	15	kroontje	kloːntjəs	kloːtsjə

2;06.06	15	krijtjes	kʀɛitjəs	kʀɛitjəs
2;06.06	15	trekt	dʀagt	dragt
2;06.06	15	trui	tʀœʏɪ	tʀœʏɪ
2;06.06	15	trap	tʀɑp	tʀɑp
2;06.06	15	uittrekken	œʏtstʀɛkə	œʏtʀɑkə
2;06.06	15	tractors	tʀɛkto:ʁɛs	tʀɑkto:ʁɛs
2;07.04	16	brandweerauto	bʀɔntʋiɹɑŏtɑŏ	bʀɔntʋiɹɑuto
2;07.04	16	brood	bʀo:tje	bʀo:tje
2;07.04	16	omdraaien	ɔmdʀa:jə	ɔmdʀa·jə
2;07.04	16	draaien	dʀə:jə	dʀa:jə
2;07.04	16	drinken	dʀɪŋk	dʀɪŋk
2;07.04	16	gedronken	xədʀɪɳk	dʀɪŋk
2;07.04	16	droog	dʀoŭχ	dʀouχ
2;07.04	16	drogen	dʀo·χə	dʀo:χə
2;07.04	16	drie	dʁi	dʁi
2;07.04	16	drie	dʀi	dʀi
2;07.04	16	drie	dʀi	dʀi
2;07.04	16	drie	dʀi	dʀi
2;07.04	16	draaimolen	dʀæimo·lə	dʀa:imo·lə
2;07.04	16	kreeg	kʀex	kʀeχ
2;07.04	16	krant	kʀɔn	kʀɔn
2;07.04	16	kroontje	klo:ntjə	klo:ntjə
2;07.04	16	kroontje	klo:ntjə	klo:ntjə
2;07.04	16	trein	tʀɛin	tʀɛin

* Broad transcriptions are identical to the CLPF transcription; narrow transcriptions are provided by the author of this thesis and focus on the representation of the cluster with greater phonetic detail.
The missing or damaged sound files are not included in the appendix.

Appendix 2: Broad and narrow transcriptions* of Robin's productions of words with /Cr/ onset clusters.

Age	session	orthography	broad transcription	narrow transcription
1;05.25	2	trein	ˈtæin̠	tɛin
1;05.25	2	trein	ˈd̪æi	tein
1;05.25	2	trein	ˈte̞in̠	tɛin
1;06.09	3	trein	ˈtɛ̞in	tɛin
1;06.09	3	trein	ˈtæ̞ĩ	tæin
1;06.22	4	trein	ˈtæ̞i·n	tɛin
1;06.22	4	trein	ˈtæ̞inə	tɛin
1;07.13	5	trein	ˈdoey̆n	tɛin
1;07.13	5	trein	ˈtɑn	tɛːin
1;07.13	5	trein	ˈṭæ̞in	tɛitɛin
1;07.27	6	trein	toe̞·n	tɛn
1;07.27	6	trein	ˈtɩ ˌty̠t ˈtoe̞ːĭn	tɛin
1;07.27	6	brood	ˈb̥ɑĭn̆t	paint
1;08.24	8	trein	ˈdɩn̠ˌdɩ̟n	dɪn
1;08.24	8	traktor	ˈtutɩ	tutɪ
1;08.24	8	brandweerauto	ˈbɑn̠·ˌto·ŭtə	pɑnto·utə
1;08.24	8	omdraaien	ˈʔoeːˈta·jə	œitajə
1;08.24	8	brandweerauto	ˈbʀɛn̠t	pʁɛnt
1;08.24	8	omdraaien	ˈʔoːˈtaĭjə	oːtaĭja
1;08.24	8	brommer	ˈbuːmə	bːomə
1;08.24	8	trein	ˈdoeyn	dʒɛin
1;08.24	8	brood	ˈbaĭt	daĭt

1;09.07	9	brandweerauto	ˈpɛn˞ˌtoŭto	pɛntoːto
1;09.07	9	brandweerauto	ˈboeːntˈboĕynt	bəɑntbœʏnt
1;09.07	9	trein	ˈdoeːn	d^{w}ɛin
1;09.07	9	trein	ˈtʊ̆ɪ̆ːn	tʊɪːn
1;09.07	9	trein	ˈdoen	
1;09.07	9	trein	ˈdɛ̝in	təɛin
1;09.07	9	trein	ˈtɛ̹in	təɛin
1;09.07	9	trein	ˈtoe̞ːy̆n	tœːːʏn
1;09.07	9	brandweerauto	ˈp^{h}ɑn̠ˌtɑuˌtɑutə	t^{h}ɑntauto
1;09.07	9	brommer	ˈbuːmɩ	buːmɪ
1;09.07	9	brandweerauto	ˈba̠ːn̠t̠	baːnt
1;09.07	9	trompet	ˈpɩt	
1;09.07	9	brandweerauto	ˈbɑn̠t	bɑint
1;09.07	9	omdraaien	ˈʔɔˈdajə	d^{w}aj
1;09.07	9	omdraaien	ˈʔoˌdajə	d^{w}aj
1;09.07	9	omdraaien	ˈʔɔmˌdajə	dʒaj
1;09.21	10	trein	ˈtoe˞n	tœin
1;09.21	10	brandweerauto	ˈta̠ːˈtoːtɔ	tant
1;09.21	10	drinken	ˈtɩŋ˞kə	dɪŋk
1;09.21	10	drinken	ˈtɩŋkə	dɪtə
1;09.21	10	trap	ʔmˈpa˞p	pəaːp
1;09.21	10	drinken	ˈd̠ɩŋkə	dɪŋkə
1;09.21	10	drinken	ˈdɩŋkə	tɪnk
1;09.21	10	trein	ʔnˈt̠ɛin	tɛin

1;09.21	10	brug	ˈpʉχ	bʉχ
1;09.21	10	omdraaien	ˌʔo̞ˈdaĭjə	odae
1;09.21	10	brandweerauto	ˈpɑn̠təˈʔot^{h}oː	ɑntəoːto
1;10.20	12	draaimolen	ˈtaĭjəˈboˑwə	dəajamoːlə
1;10.20	12	brandweerauto	ˈpɑːn̤təˈʔoŭto	pɑantəoːto
1;10.20	12	brandweer	ˈɦɑntəˈfɹ̟	hantəfi
1;10.20	12	broertjes	ˈbuˑĭ	bun
1;10.20	12	drinken	ˈtɩŋkə	tɪŋkə
1;10.20	12	drinken	ˈt^{h}ɩŋkχ̆ə	tsɪiŋkə
1;10.20	12	drinken	ˈtɩŋkə] [ˈʔɔpɩn]	d^{w}ɪŋkə
1;11.06	13	trein	ˈtoeːĭn	t^{h}ɛin
1;11.06	13	trein	ˈtoe̞ːin	tʁɛin
1;11.06	13	trein	ˈta̞ˑn	tain
1;11.06	13	draaimolen	ˈtaːˌmol̠ə] [ˈnɑu]	t^{h}aːnmolə
1;11.06	13	draaimolen	ˈtaˌmoːɬ̠ə	t^{h}aːimo
1;11.06	13	draaimolen	ˈtaːˌmoːʎ	taːnmolə
1;11.06	13	draaimolen	ˈtaˑˌmoːj̆ə	taːmojə
1;11.06	13	draaien	ˈtɛ̠ijə	taːjə
1;11.06	13	drinken	ˈt^{h}ɩːkə	t^{h}ɪŋkə
1;11.06	13	brandweer	ˈba̠n̤t̤ˌʋe̹ɹ	bɑntʋeɹ
1;11.06	13	brandweer	ˈp^{h}a̠n̤tˈfʌːɹ	p^{h}ɑntʋiɹ
1;11.06	13	brandweer	ˈbatˌfëːɹ	p^{h}ɑntʋiɹ
1;11.06	13	straat	ˈtaːt	taːiht
1;11.06	13	straat	taːt	tɑaːt
1;11.06	13	trekt	tɪk	tɪk

1;11.06	13	kruk	tɪk	tsɪk
1;11.06	13	brandweer	ʔəbãːfe	pantəʋeɹ
1;11.20	14	trein	ˈtɛˑn	tɛin
1;11.20	14	trein	ˈtɛ̝ĭn	tɛin
1;11.20	14	treinen	ˈtɛin̠ə	tɛin
1;11.20	14	brood	ˈboːt	bʷoːt
1;11.20	14	brandweerauto	ˈbɑt̠əˌʋe̠ʴˈʔaŭt̠ä	bɑtɪʋɛɹ
2;00.04	15	trein	ˈtæɦ̆in	tɛin
2;00.04	15	drinken	ˈtĩːkə	tɪkə
2;00.04	15	draaimolen	ˈdɑt̠ˈmol̠ə	dəaːtmolə
2;00.04	15	draaimolen	ˈdaĭˌmol̠ə	daːimolə
2;00.04	15	draaimolen	ˈtaiˈmoːɫ̠ə	taimolə
2;00.04	15	draaimolen	ˌtaˈmoːl̠ə ˈʔɯn	tləamoːlə
2;00.04	15	draaimolen	ˈtæ̝iˈmol̠ə	taimoːlə
2;00.04	15	draaimolen	ˈtaĭˈmoːl̠ə	tlɛinmolə
2;00.18	16	krijgen	ˈt͜sɛiχə	tsʷɛiχə
2;00.18	16	brandweerauto	ˈʋăt̠ˌʋe̠ʴˈʔoŭto	ʋəɑntʋeɹ
2;00.18	16	brommer	ˈbɔməˑɹ	bɔmə
2;00.18	16	drinken	ˈdɯŋkə	dɪŋkə
2;01.06	17	tram	ˈpɩˑm	pɪm
2;01.06	17	trein1	ˈtɛ̝in	tɛin
2;01.06	17	trein2	ˈtɛ̝in	tɛin
2;01.06	17	brandweerauto	ˈbɑ̝n̠tˌʋe̠əˈʔɑ͜oto	bɑnt
2;01.06	17	brandweerauto	ˈbɑn̠ˑˈdʋeʴˈʔɑ͜otoˑ	bəɑnt

2;01.06	17	draaimolen	ˈdæ̞iˈmoˑˌl̠ə	daːimoːlə
2;01.06	17	trein	ˈtɛ̞in	tɛin
2;01.06	17	drukt	ˈḍʉkt	dəɪk
2;01.06	17	trekken	ˈtɛ̠ˑkə	
2;01.06	17	duplotrein	ˈpeːpoˑs̠ˈt͜s̠ɛin	tsɛin
2;01.06	17	drinken	ˈdɯŋkə	dʷɪŋkə
2;01.06	17	drinken	ˈlɯŋkə	tʃɪŋkə
2;01.06	17	drinken	ˈdɯŋk	dɛŋχ
2;01.25	18	treinen	ˈtʰɛinə	thɛinə
2;01.25	18	zonnebril	ˈsɔ̞ɬˈbɛˑu	bɛœʏu
2;01.25	18	zonnebril	ˈs̠ɔndəˈbɯ	bɪ
2;01.25	18	springen	ˈspɯnə	spɪnə
2;01.25	18	draaimolen	ˈda̝imu̞ːs	daːimoːlə
2;01.25	18	draaimolen	ˈdæ̃iˈmoːlə	djaimoːlə
2;01.25	18	draaimolen	ˈtaiˈmoijət͜jə	tjaimoːlə
2;01.25	18	trein	ˈtʰɛin	tsəɛin
2;01.25	18	brandweerauto	ˈbɑntʋɪ̝əˈʔoŭto̞	bɑntʋeɹ
2;01.25	18	brandweerauto	ˈbɑntʋɪ̝əˈʔoŭto̞	bɑntʋeɹ
2;01.25	18	draaien	ˈd̠aˑĭjə	dʰɛinən
2;01.25	18	draaimolen	ˈdʰæiˌmoːlə	djaimoːlə
2;01.25	18	drinken	ˈdɯŋkə	dɪŋkə
2;01.25	18	trein	ˈtʰɛ̞in	tʰəɛin
2;01.25	18	trein	ˈtʰɛ̞in	teɛin
2;01.25	18	draaimolen	ˈdaˑiˈmoːlə	daːimoːlə
2;01.25	18	draaimolen	ˈdæiˌmoːl̠ə	daːimoːlə

2;01.25	18	krijgt	ˈtʰɛĭk	kjɛiχ
2;01.25	18	draaimolen	ˈtæiˈmoːŭl̠ə	dlœʏmoːlə
2;01.25	18	draaimolen	ˈda̠iˈmoŭl̠ə	daːimoːlə
2;01.25	18	brand	ˈbɑ̟n̠t	bəɑnt
2;01.25	18	brandweer	ˈbɑ̟̃ntˈʋeʴ	bəɑntʋeɹ
2;02.27	19	trein1	ˈtʰɛin	tʰɛin
2;02.27	19	brandweerauto	ˈbɑːm̆ʋeʴˈʔaŭtoːs̠	bəɑntʋeɹ
2;02.27	19	traktor	ˈtɑkˈtɔʴ	tɑktɔɹ
2;02.27	19	brandweer	ˈbɑn̠ˑˈdveːɹ	bəɑntʋeɹ
2;02.27	19	hijskraan	ˈɦɛiʃ̟kaˑn	kaːn
2;02.27	19	hijskraan	ˌhɛis̠ˈkan	kaːn
2;02.27	19	trein	ˈtɛin	tɛin
2;02.27	19	zonnebril	ˈs̠ɔˑnəˈbɯ	beɪu
2;02.27	19	drinken	ˈdɯŋkə	dɪŋkə
2;02.27	19	drinken	ˈdɯŋkə	dɪŋkə
2;03.10	20	draaimolen	ˈdaiˌmoˑlə	dlaːimoːlə
2;03.10	20	trein	ˈtʰæin	thɛin
2;03.10	20	trekken	ˈtˑɩkə	tjikə
2;03.10	20	broodje	ˈboˑt͜jə	bwoːtjə
2;03.10	20	broodje	ˈboːt͜jə	boːtjə
2;03.10	20	broodje	ˈboˑt͜jə	bəoːtjə
2;03.10	20	broodje	ˈbʁ̣̆oːt͜jə	bʁoːtjə
2;03.22	21	brood	ˈḅoːt	bəoːt
2;03.22	21	straks	ˈtʰɑks	tʰɑks

2;03.22	21	drinken	ˈdɩŋkə	dɪiŋkə
2;03.22	21	drinken	ˈdɩŋkə	diːŋkə
2;03.22	21	broodje	ˈbozə	bɛ oːt
2;03.22	21	dromedaris	ˈdɔməˈdaːɫəs	dəɔmə
2;03.22	21	dromedaris	ˈdɔːməˈdaʀ̣̆əs̠	dwɔːmə
2;03.22	21	krokodil	ˈkoˑkəˈdɩɔ	koːkədɪu
2;03.22	21	krokodil	ˈkoːkəˈbɩu	koːukɔdɪu
2;03.22	21	brandweerauto	ˈbɑn̠ˑˌʋɩˈotʰo	bəɑntʋɪoto
2;03.22	21	drinken	ˈdɩŋkə	dɪŋkə
2;03.22	21	hijskraan	ˈɦɛisˌkan	təɑn
2;03.22	21	hijskraan	ˈɦɛ̝isˌkan	kaːn
2;03.22	21	brommer	ˈbumaˀ	bumə
2;03.22	21	trein	ˈtɛ̝in	tsɛin
2;03.22	21	drinken	ˈdɩŋkə	dɛɪŋk
2;04.07	22	brandweerauto	ˈbɑ̃ň̆ʋɩˈoːto	bəɑntʋɪoto
2;04.07	22	trein	ˈtɛin	təɛin
2;04.07	22	trein	ˈtʀɛin	tɛin
2;04.07	22	brandweer	ˈbɑntˈʋe̠ɹ	bɑntʋeɹ
2;04.07	22	brandweerazerne	ˈbɑṇtˈʋɩɹˈʋɑsɑˈsɛˑ	ʋɑntʋu
2;04.07	22	kroontje	ˈkɑ͜ũntj͜ə	kəamtjə
2;04.07	22	kroon	ˈkɑ͜ũn	kan
2;04.07	22	kroontje	ˈkɑ͜ontj͜ɩ	kɑoːntj
2;04.07	22	prinsje	ˈpɩn̠ʃ̟ə	pɪinʃə
2;04.07	22	prinsje	ˈpʌ̃ːtj͜ə	pəɪntjə
2;04.07	22	trein	ˈtʰoe̦yn	tsœʏn

2;04.07	22	trommel	tɔməɫ	tʁɔmɔm
2;04.07	22	trommel	tɔməɫ	tjɔməɫ

* Broad transcriptions are identical to the CLPF transcription; narrow transcriptions are provided by the author of this thesis and focus on the representation of the cluster with greater phonetic detail.

** The numbers following the item, refer to the order of appearance in the CLPF database.

*** The missing or damaged sound files are not included in the appendix

Appendix 3: Broad and narrow transcriptions* of Tirza's productions of words with /Cr/ onset clusters.

age	session	orthography	broad transcription	narrow transcription
1;08.05	2	drie	ˈdʂ̠̆i	tli
1;08.05	2	trein	ˈtʀ̣ɛinə	təɛinə
1;08.05	2	trein	ˈɦfɛinəː	hɛinə
1;10.08	5	broek	ˈpuːk	pʉuk
1;10.22	6	kruiwagen	ˈhɑuˈtǎ̠χˑə	təhauf
1;10.22	6	krauw	ˈt͜jau	tjau
1;11.08	7	broek	ˈbuˑk	b^{w}uk
1;11.19	8	traktor	ˈṭǎ̠k	t^{j}ak
1;11.19	8	traktor	ˈt̠ɑk	t^{j}ak
1;11.19	8	traktor	ˈdɑk	dɑk
1;11.19	8	trein	ˈdɛin	dɛin
1;11.19	8	pruimen	ˈtoeymˑə	t^{h}aœʏmə
1;11.19	8	broek	ˈbʀ̠̆uk	bwuk
2;00.05	9	bruine	ˈboe̜ynəː	bəaunə
2;00.05	9	krokodil	ˈtoe͜ˑyːtəˌto̜̝	tɑutə
2;00.05	9	krokodil	ˈtɔːtəˌtɯ	təoːtə
2;00.05	9	draaimolen	ˈtaiɩˈmoˑlə	tajəmoːlə
2;00.05	9	draai	ˈtai	tai
2;00.05	9	draaien	ˈtaˑiə	t^{h}aiə
2;00.05	9	ronddraaien	ˈhɔ̜̆ntˈtaiə	hɔnt
2;00.05	9	drinken	ˈlinˌkə	lʒɪnkə
2;00.05	9	drenthe	ˈlɛnˑtə	l^{j}ɛntjə

2;00.05	9	hijskraan	ˈʔɛ̯iˈtaːn	kʲaːn
2;00.05	9	drenthe	ˈlɛntə	tɛntə
2;00.18	10	draaimolen	ˈdaˑiəˈmo̜ːl̠ə	daːjəmoːlə
2;00.18	10	draaimolen	ˈdaiˌmoːl̠ə	dəaimoːlə
2;00.18	10	draaimolen	ˈd̠aiˌmoːlə	daimoːlə
2;00.18	10	draaimolen	ˈdaiˌmoːlə	daimoːlə
2;00.18	10	krokodil	ˈkaːkəpə̆ˈt̠oˑu	kaɛəkədou
2;00.18	10	druiven	ˈɦa̝ːfə	hœʏfə
2;00.18	10	druiven	ˈɦoeːyfə	hœʏfə
2;00.18	10	draaimolen	ˈdɛiˌmoːlə	djɛimoːlə
2;00.18	10	kraan	ˈtaːn	tʲain
2;01.02	11	draaimolen	ˈdɛiˌmoːl̠ə	dəɛiimoːlə
2;01.02	11	draaimolen	ˈdeimoˑˌləː	deimoːlə
2;01.02	11	draaimolen	ˈdeːmoŭl̠ə	deimoːlə
2;01.02	11	draaimolen	ˌd̠ɑuˈmoŭˑɫə	djaɑumoulə
2;01.02	11	draaimolen	ˈtɛiˌmoːlə	tɛimoːlə
2;01.02	11	brandweerauto	ˈbæ̠̆ntəˌʔoeuto	bænt
2;01.02	11	kraan	ˈtaːn̠ə̆	taːnə
2;01.02	11	trein	ˈdɛ͜iˑn	djɛin
2;01.02	11	afdrogen	ʔəˈχoeːuxə	əχəouχə
2;01.02	11	drinken	ˈdɩnˈkə	dʲɪŋkə
2;01.02	11	draaien	ˈɦæ͜ijə	ɦæijə
2;01.17	12	draaimolen	ʔnˈdeĭmoːl̠ə	deimoːlə
2;01.17	12	draaimolen	ˈdɛimoːl̠ɛ	dəaimoːlə

2;01.17	12	draaimolen	ˈtaiˑˌmoˑlə	təaimolə
2;01.17	12	krokodil	ˈkoŭˈkə̩ɫ̌ˌtɔ	koukəlto
2;01.17	12	draaimolen	ˈdɛˑimoːˌlə	tɛimoːlə
2;01.17	12	krokodil	ˈkɑ͜okəˈdœː	kɑoːkə
2;01.17	12	drie	ˈdl̩i	dəli
2;02.00	13	draaimolen	ˈtɛimoːˌlə	dɛːmoːlə
2;02.00	13	draaimolen	ˈtaiˑmˌmoːlë	tɛimoːlə
2;02.00	13	draaimolen	ˈta̠imoˑlə	təaimoːlə
2;02.00	13	krokodil	ˈkøˑkədø̠̆ˑ	koːkəkiu
2;02.00	13	krokodil	ˈkoeˑykəˈdaː	kəœukə
2;02.00	13	trui	ˈt̠lay̨	klaʏ
2;02.00	13	kraan	ˈtaːn̠	taːnə
2;02.00	13	krijgt	ˈtɛ̩iχ	tɛis
2;02.00	13	krijgt	ˈtɛ̩iχ	tɛiχ
2;02.00	13	trommel	ˈχɔˑmɑ̆χ̌	χɔmɑχ
2;02.12	14	draaimolen	ˈtʀaˑiˈmoˑlə	təʁaimoːlə
2;02.12	14	trein	ˈtɛ̩inə	tɛin
2;02.12	14	draaimolen	ˈtaimoˑl̠ə	təaimoːlə
2;02.12	14	krijgen	ˈtɛiχəː	tɛiχə
2;02.12	14	draaimolen	ˈtail̩əˈmoːlə	tailə
2;02.12	14	draaimolen	ˈdaimoˑlə	dəajemoːlə
2;02.12	14	draaimolen	ˈd̪̠aiˌmoːlə	daimoːlə
2;02.12	14	draaimolen	ˈdeiˌmoːlə	daimoːlə
2;02.12	14	draaimolen	ˈtɛ̠iˌmolə	teimoːlə
2;02.12	14	trommels	ˈtlʌməɹs	tlʌməɹs

2;02.12	14	trommelaar	ˈtɔməlaˀ	tɔməlaː
2;02.12	14	draaimolen	ˈtɛ̆iˌmolɩˌɕə	tɛimolɪtjə
2;02.25	15	draaimolen	ˈdæˑiˈmoːl̠ə	dəaimoːlə
2;02.25	15	trein	ʔəˈɣɛiˑn	χəχɛin
2;02.25	15	trein	tə̂ʀɛin	təʀɛin
2;02.25	15	broek	ˈbʁuk	bʁuk
2;02.25	15	draaimolen	ˈdæiəˈmoːɩjələ	dainəmoːlə
2;02.25	15	draaimolen	ˈχaiˌmoˈmoːlə	χaimomoːlə
2;02.25	15	draaimolen	ˈtaiˌmoːlə	tʲamoːlə
2;02.25	15	drinken	ˈdʁɪ̯ŋˑkə	dəʀɪŋkə
2;03.12	16	draaimolen	təˈʀaiˌmoːlə	dəχədjemoːlə
2;03.12	16	draaimolen	də̂ˈʀaiˌmoˑlə	təʀajəmoːlə
2;03.12	16	draaimolen	dəˈʀaĭjəˌnɑ͜olə	kʀaimoːlə
2;03.12	16	draaimolen	də̂ˈʀaiˌmoˑlə	təʀɛimolə
2;03.12	16	broek	ˈbʀˑuk	bwuk
2;03.12	16	kraan	ˌkəˈʀˑaˑn̠	təʀaːn
2;03.12	16	hijskraan	ˈʔɛ̝isˌkəʀaˌhaˑn	təʀəaːn
2;03.12	16	trein	ˌdə̂ˈʀːɛin	təʀɛin
2;03.12	16	trap	təˈʁɑp	təʁɑp
2;03.27	17	draaimolen	ˈtʀɛiˈmoˑlə	tʀɛimoːlə
2;03.27	17	draaimolen	ˈtʀæiˈjəˈmoˑlə̝ˑ	tsʀɛimoːlə
2;03.27	17	draaimolen	ˈdʀæ̝iˈjəˈmoˈlʌ̈	dʀɛimoːlʌ
2;03.27	17	strooien	ˈtʀoˑjə	χoːjə
2;03.27	17	strooien	sˑˈʀoˑjə	χoːjə

2;03.27	17	strooien	χoːjə	χoːjə
2;03.27	17	trein	ˈtʼʀɛin	tʀɛin
2;03.27	17	trein	ˈʂ̠̌ɔ̜ɹʂ̠̌tə̂ˌʁɛin	thɛin
2;03.27	17	trein	ˈtʁɛin	tʀɛin
2;05.05	18	druiven	ˈdʀoeyfə	dʀœʏfə
2;05.05	18	druiven	ˈdʀoe̩yfə	dʀœʏfə
2;05.05	18	draaimolen	ˈdʀaiˌmo˙lï	dʀaːjmolə
2;05.05	18	draaien	dʀaːi	dʰʀaiː
2;05.05	18	draaimolen	ˈθʀaːˌmo˙lə	tsaimolə
2;05.05	18	brommer	ˈbʀɔmə	bʁəɔmə
2;05.05	18	drie	ˈtə̆ʀi	təʀi
2;05.05	18	drie	ˈdʀi	tχi
2;05.05	18	trein	ˈtʀɛ̩in	tʀɛi
2;05.05	18	trein	ˈtʀɛ̩in	tʀɛin
2;05.05	18	draaien	ˈdʀa˙iə	dʀaːjə
2;05.21	19	tracteren	ˌtɑkˈteːʀə	tɑk
2;05.21	19	trommel	ˈtʀɔm˙ʌ	tʀɔmʌ
2;05.21	19	trein	ˈtʀɛin	tʀɛin
2;05.21	19	trekken	ˈtʀʉkə	tʀɛkə

* Broad transcriptions are identical to the CLPF transcription; narrow transcriptions are provided by the author of this thesis and focus on the representation of the cluster with greater phonetic detail.
The missing or damaged sound files are not included in the appendix.

Appendix 4: Broad and narrow transcriptions* of Enzo's productions of words with /Cr/ onset clusters.

age	session	orthography	broad transcription	narrow transcription
1;11.08	1	afbreken	ˈʔɑmʀet̠s̠	ɑndʁeːts
1;11.08	1	afbreken	ˈʔɑmʀe	ɑbʀeː
1;11.08	1	broek	ˈbʀ̣uk	bruk
1;11.08	1	bramen	paˈlaˑmə	pəlaumə
1;11.08	1	brandweerauto	ˈbʀɑn̠ˌt̠oŭt̠o	bʀɑnt
1;11.08	1	trein	tə̆ˈʀɛin	tʀɛinə
1;11.08	1	trein	tə̆ˈʀɛin	təʀɛin
1;11.08	1	traktor	ˈtɑktɔ	təɑktɔ
1;11.08	1	traktor	tə̆ˈʀɑktɔ	təʀɑktɔ
1;11.08	1	traktor	ˈtʀɑːktɔ	təʀɑktɔ
1;11.08	1	traktor	ˈtʁɑktɔ	təʀɑktɔ
1;11.15	2	aangetrokken	tʀɛkt	tʁɛts
1;11.15	2	brandweerauto	ˈbʀɑn̠ˑˌt̠otoˑ	bʀɑnt oto
1;11.15	2	brandweerauto	ˈblɑ̟n̠ˌdoŭto	blɑnt oto
1;11.15	2	drie	nəˈʀi	nuʁi
1;11.15	2	drie	ˈdʀi	dwi
1;11.15	2	traktor	ˈtʀ̣ɒkˌtɔ	tʷɑktɔ
1;11.15	2	traktor	tə̆ˈʀɑktɔ	təʀɑktɔ
1;11.15	2	trap	tə̆ˈʀɑ	təʀɑk
1;11.15	2	trein	ˈʀ̆ɛin̠	ʀɛin
1;11.15	2	brandweerauto	ˈbʀɑn̠ˌt̠oŭt̠o	bʀɑnt
1;11.15	2	trein	təˈʀɛin̠	tɛrɛin
1;11.29	3	trein	ˈtʀĕ	tʀe
1;11.29	3	bril	ˈbʀɪˑ	bʀɪ
1;11.29	3	trein	ˈdʀe	dʁeː
1;11.29	3	tekken	ˈtʀɛkɛ̟	tʀɛkɛ
1;11.29	3	trein	ˈtʀe̟in	tʀɛin
1;11.29	3	omdraaien	ˈdʀaˑjə	omdʀaːjə

1;11.29	3	omdraaien	ˌdʀaijə	umaʀaːjə
1;11.29	3	trein	ˈtʀeː	tʀeːn
1;11.29	3	trein	ˈtʀeː	tʁeː
2;00.13	4	draaimolen	ˈdʀ̩aˑˌmoˑlə	dəʀa̶imoːlə
2;00.13	4	brengt	ˈbʀ̩ɛŋt	bəʀɛnt
2;00.13	4	traktor	tˈʀ̩ɑktɔ	tʀɑktɔ
2;00.13	4	trekken	ˈtʀɛkɩ	tʀɛkə
2;00.13	4	krokodil	kʁɔ̆kɔ̆ˈdɩɔ	kukɔdiɫ
2;00.13	4	drinken	ˈdʀ̩ɩŋːkə	dʁɪŋkə
2;00.13	4	drinken	ˈd̪oe̱ŋkə	dʉɳk
2;00.13	4	opdrinken	ˈɔ̆mˌdʀɪ̈ŋkə	dəʀɪŋkə
2;00.13	4	trein	ˈtʁeˑn̩	tʀɛin
2;01.03	6	druppeltjes	ˈdʀʉpəç̟ət	drʉp
2;01.03	6	citroen	tiˈtʀuň	tɪtʀun
2;01.03	6	streepjes	ˈtʀeˑpjɛ̈t	tʀepjɛt
2;01.03	6	citroenen	tiˈtʀunə	titrunə
2;01.03	6	krokodil	ˌkʁ̆ɔ̆kə̥ˈdɩɔːl̆	kʰɔk
2;01.03	6	krokodil	ˌko̞koˈdɩl	kɔk
2;01.03	6	trein	ˈtʀɛiň	tʀɛin
2;01.17	7	gebraden	ˇbʀaˑt	bɔʀɑts
2;01.17	7	drinken	ˈdʀ̩ɩŋːkə	dʀɪŋkə
2;01.17	7	broertje	ˈbʀuːcə̆	bʀʉtjə
2;01.17	7	brandweerman	ˈbʀɑˑˌmɑm	bʀɑmɑm
2;01.29	8	drie	ˈdʀi	tʀi
2;01.29	8	broek	ˈbʀuk	bʀuk
2;01.29	8	krokodil	kʁːˌkɔ̈ˈdɩ̠l	kʀəkɔdɪɫ
2;01.29	8	traktor	ˈtʀɑʔˌkɤ̈	tʀɑktɔ
2;02.14	9	aantrekken	ˈʔanˈtʀɛkə	tʀɛkə
2;02.14	9	brandweerauto	ˈloeyly	plœʏlʏ
2;02.14	9	brandweerauto	ˈtʀɑntɩˌʋɩˈotoŭ	tʀɑnt
2;02.14	9	brandweerauto	ˈpʁɑntəˈmɑn	pʁɑntə
2;02.14	9	brug	ˈbuɪ̆t	put
2;02.14	9	brand	ˈbʀɑ̟n̠t	bʀɑnt

2;02.14	9	brandweerman	ˈbʀantəˌʋɩˈmãn	bʀantə
2;02.14	9	trein1	ˈtʀɛin	tʀɛin
2;02.14	9	krokodil	ˈnɔkɔˌdɪɫ	mɔkɔdɪl
2;02.14	9	traktor	ˈniˈtʀɑktɔ	tʀɑktɔɹ
2;02.14	9	opdraaien	ˈʔɔpˈdʀaːʝ̆ə	ɔpdʀaːjə
2;02.14	9	driewieler	ˈtʁ̣iʋiˌlɛ̈	tʁi
2;02.14	9	drie	ˈdʀi	dʀi
2;02.14	9	hijskraan	ˈɦɛiˌkʀan	hɛikʀaːn
2;02.14	9	traktor	ˈtʀɑktɔ	təʀɑktɔɹ
2;02.14	9	traktor	ˈtʀɑktɔ	tʀɑktɔ
2;02.14	9	trui	təʀ̆oẹːːu	təʀœːu
2;02.14	9	drie	ˈdʀi	dʀi
2;02.14	9	aantrekken	ˈʔa·nˌtʀɛkə	aːntʀɛkə
2;02.14	9	drinken	ˈdʀɩŋ·kχə	dʀɪŋkə
2;02.14	9	trein	ˈtʀɛin	tʀɛin
2;02.14	9	krokodil	ˌkʁokəˈdɪɫ	tʁœʏfuto
2;02.14	9	struisvogel	ˈtʀoeyˌfu·t͜jọ	tʀœʏɫ
2;02.14	9	trui	ˈtʀoe͜ːu̯	tʀœʏɫ
2;03.14	10	brandweerauto	ˈbʀã·tʼʋɩˈʔot̪o	bʀatʋɪʔoto
2;03.14	10	brandweermanstuur	ˈʀamɑˌmɑnˈty	rːamɑ
2;03.14	10	traktor	ˈtʀɑktɔ	tʀɑktɔ
2;03.14	10	brand	ˈbʀãnt̪	bʀɑnt
2;03.14	10	traktor	ˈtʀɑḳˌt̪ɔ	tʀɑktə
2;03.14	10	krokodil	ˈkukü'dɩǔ	kʁukudiɫ
2;03.14	10	brandweerhoed	ˈloeyˌlu·t	tlalut
2;03.14	10	drinken	dəˈʀɩŋ	dəʀɪŋk
2;03.14	10	drie	ˈdʀĭ	dʀi
2;03.14	10	traktor	ˈtʀ̣ɑḳtɔ	tʀɑktɔ
2;03.14	10	traktor	ˈtʀɑktɔ	sʀɑktɔ
2;03.28	11	brandweerauto	ˈpʀɑṇʋɩˈʔot̪o	pʀɑn
2;03.28	11	citroen	ˌsiˈtʁũn	sitʁun
2;03.28	11	hijskraan	ˈɛiˌkʀa·n	kʁaːn

2;03.28	11	drinken	ˈdʀɩŋkə	dʀɪŋkə
2;03.28	11	hijskraan	ˈɦɛiˌkʀaṉ	kʁaːn
2;03.28	11	gebroken	ˈbʀokə	bʀoːkə
2;03.28	11	brug	ˈbʀ̣ʉχ	brʉχ
2;03.28	11	drink	ˈdʀɩŋk	tʀɪŋk
2;03.28	11	brandweerauto	ˈloeˑly̌ˈʔoto	laːləoto
2;03.28	11	trompet	ˌdʀɔmˈpɛt	tʀɔmpɛt
2;03.28	11	traktor	ˈtʀɑkt̠ɔ	tʀɑktɔɹ
2;03.28	11	drukken	ˈdʁɩ̧kə	dʀɪkə
2;03.28	11	drukken	ˈdʁʉkə	dʀukə
2;03.28	11	hijskraan	ˈʀɛiˈkʀaˑṉ	ʀɛikʀaːn
2;03.28	11	hijskraan	ˈɦɛ̠iˈkʀaˑn	kʀaːn
2;03.28	11	hijskraan	ˈʀɛiˌkʀaˑn	ʀɛikʁaːn
2;03.28	11	hijskraan	ˈhɛiˌkʀan	kʀaːn
2;03.28	11	traktor	ˈtʀɑkˈtɔ	tʀɑktɔ
2;03.28	11	trompet	tŭmˈpæ	tumpa
2;03.28	11	hijskraan	ˈɦɛiˈkʀa̠ṉ	kʁaːn
2;04.11	12	prik	ˈpʀɩk	pʀɪk
2;04.11	12	brand	ˈbʀɑnˑ	bʀɑn
2;04.11	12	brand	ˈbʀɑn	bʀan
2;04.11	12	traktor	ˈtʀɑktɔʁ	tʀɑktɔɹ
2;04.11	12	breekt	ˈbʀeːĭt	bʀeːt
2;04.11	12	prink	ˈpʁɩk	pʁɪk
2;04.11	12	traktor	ˈtə̂ʀɑktɔʁ	tʀɑktɔɹ
2;04.11	12	drie	ˈdʀɩ	tʀi
2;04.11	12	brandweer	ˈbʀɑ̃nˌʋeɹ	ˈbʀɑ̃nˌʋeɹ
2;04.11	12	prikje	ˈfʀɩk˺jə	kfʀɪk
2;04.11	12	breekt	ˈbʀeĭt	pʁiːt
2;04.11	12	broek	ˈbʀuk	bʀuk
2;04.11	12	springhanen	ˈpʀɩŋkˈhaˑnə	pʀɪŋk
2;04.25	13	trap	ˈtə̂ʀɑp	tʀɑp
2;04.25	13	trap	tˈʀɑp	tʀap

2;04.25	13	brandweermannen	ˈbʀɑn˙ˌtʋɩˈmɑn̠ə	bʁɑnt
2;04.25	13	traktor	ˈtʀɑkˌtɔʁ̆	tʀɑktɔɹ
2;04.25	13	traktor	ˈtʀɑkˌtɔ	tʀɑktɔ
2;04.25	13	treinen	ˈtʁɛinə	tʁɛin
2;04.25	13	brandweermannen	ˈbʀɑ̃n̦ʋɩˈmɑnə	bʀɑnt
2;04.25	13	brandweerhoed	ˈbʀɑn̠dəˈɦut	bʀɑnd
2;04.25	13	bril	ˈbʀɨ˙ʋ	bʀɪɫ
2;04.25	13	treinen	ˈtə̂ʀɛin	təʀɛin
2;04.25	13	brandweerhoeden	ˈbʀɑndˈˈɦu˙tə	bʀɑnt
2;04.25	13	drie	ˈd̪̆ʀi	dʀi
2;05.09	14	krant	ˈkʀɑn	kʀɑn
2;05.09	14	brandwerauto	ˈbʀɑ̃mˈʔo˙to˙	bʀɑm
2;05.09	14	brandweermannen	ˈbʀɑ̃n˙tˈʋe̠mɑ̥n̥ə̥	bʀɑn
2;05.09	14	brandweerhoed	lɛːlut	lœʏlut
2;05.09	14	trompet	ˌtʰɔmˈpʰɛt̪	thɔmpɛt
2;05.09	14	draaimolen	ˈdʀaiˌm̌olɛ̈	dʀaːimolɛ
2;05.09	14	brandweerhoed	ˈbʀɑ̃nˌʋ̆eˈhut	bʀɑntʋɪhut
2;05.28	15	traktor	ˈtʀɑktɔ	təʀɑktɔɹ
2;05.28	15	traktor	ˈtʀɑktɔ	tʀɑktɔ
2;05.28	15	brommers	ˈbʀɔmɛ̈s̪	bʁɔmɛs
2;05.28	15	brandnetels	ˈbʀɑ̃n̦ˈʔnetɔ̈s̪	bʀɑnt
2;05.28	15	brommer	ˈbʀɔməʁ	bʀɔməχ
2;05.28	15	trein	ˈtʀɛ̟in	tʀɛin
2;05.28	15	kraaltjes	ˈkʀɑu˙ç̟ɩs̪	kʀɑutjəs

2;05.28	15	traktor	ˈtʀ̣ɑktɔ˙	tʀɑktɔ
2;05.28	15	traktor	ˈtʁɑktɔ	tʀɑktɔ
2;06.11	16	brommer	ˈbʀɔm	bʀɔm
2;06.11	16	brand	ˈbʁɑ̰̃nt	bʀɑnt
2;06.11	16	brand	ˈbʀɑ̃ː	bʁɑn
2;06.11	16	brandweer	ˈloeyly	lalʏ
2;06.11	16	draaien	ˈdʀa˙iə	dʀaːjə
2;06.11	16	brommer	ˈbʀɔməʁ̆	bʀɔməɹ
2;06.11	16	trapje	ˈtʁ̆ɑpç	tʀɑpjə
2;06.11	16	brand	ˈbʀɑ̃nt	bʁɑnt
2;06.11	16	brand	ˈbʀɑn	bwɔntə
2;06.11	16	brandweer	də̆ˈloeːly	dlœdʏ
2;06.11	16	drum	ˈdʀʉm	dʀʉm
2;06.11	16	trompet	ˌtʀɔmpɛ̰˙t	tʀɔmpɛt
2;06.11	16	brandweerpakken	ˈloeːly̜t ˈpɑkə	dladʏt
2;06.11	16	brandweerman	bʀɑ̃nˌʋɩ̢mɑn̠	bʀɑnʋɪmɑn
2;06.11	16	traktor	ˈtʀɑktɔʁ̆	tʀɑp
2;06.11	16	traktor	ˈtʀɑktɔʁ̆	tʀɑktɔɹ
2;06.11	16	brandweermannen	ˈbʀɑ̃nˌtʋɩˈmɑnə	bʀɑnt
2;06.11	16	drinken	ˈdʀɩŋkə	dʀɪŋkə
2;06.11	16	brommer	ˈbʁɔməʁ	bʁɔmə
2;06.11	16	brandweermannen	ˈbʀɑnˌʋɩˈmɑn̠ə	bʀɑn
2;06.11	16	bril	ˈbʀɩʊ̆	bʀɪɫ
2;06.11	16	draaien	ˈdʀaiə̆	dʀaːjə
2;06.11	16	brandweerman	ˈbʀɑnˈʋɩ̢mɑn	bʀɑn

2;06.11	16	brandweermannen	ˈbʀɑntɩˌmɑnə	bʀɑnti
2;06.11	16	brandweerauto's	ˈloey̌lyˈʔo̧toːs	lalɤoto
2;06.11	16	propellers	ˌpuˈpɛ̞nɩs̪	pupɛnɪs

* Broad transcriptions are identical to the CLPF transcription; narrow transcriptions are provided by the author of this thesis and focus on the representation of the cluster with greater phonetic detail.

The missing or damaged sound files are not included in the appendix.

Appendix 5: Broad and narrow transcriptions* of Eva's productions of words with /Cr/ onset clusters.

age	session	orthography	broad transcription	narrow transcription
1;04.12	1	brood	ˈmo˞p	mop
1;04.12	1	brood	ˈmo˞p	mop
1;04.12	1	prik	ˈt̺ɩt̺	tɪt
1;04.12	1	prik	ˈtɩt	tɪt
1;04.12	1	prik	ˈȟɩt	hɪt
1;04.12	1	prikken	ˈte˞ĭtə	teitə
1;05.22	2	brood	ˈboːp	bəoːp
1;05.22	2	prik	ˈt̺ɩt̺	tɪt
1;05.22	2	drinken	ˈhɩːn̠ɛ̈	dɪːnɛ
1;06.01	4	trein	tɛ̞in	tɛin
1;06.01	4	trui	ˈtoey	tɛœʏ
1;06.01	4	broek	ˈbut	but
1;06.01	4	kraanwagen	ˈtaŋˈlaːχə	taːŋ
1;06.01	4	brood	ˈboŭt̺	boːt
1;06.11	5	draaien	ˈɦæiĭ̆jə	hæijə
1;06.11	5	draaien	ˈdaiĭ̆jə	daːjə
1;06.11	5	brilletje	ˈpɩʊɛ̈	pɪʊə
1;06.11	5	draaien	ˈt˥a̞ijə	dəaːjə
1;06.11	5	drie	ˈti	diː
1;07.15	6	trein	ˈtɛ̠in	tɛin
1;07.15	6	krant	ˈtɑn̠	təɑn
1;07.22	7	broek	ˈbuːt	bʉut
1;08.12	8	citroenen	ˈtunə	tunə
1;08.12	8	trommel	ˈtɔm˞ɔ	twumɔ
1;08.12	8	trommel	ˈtɔmɔ	tʷɔməɫ
1;09.08	9	krokodil	ˈd̺̟ɩo	
1;09.08	9	krokodil	ˈtotoˈdɩɔ˥	toto
1;09.08	9	krokodil	ˈdɩɫ̠	
1;09.22	10	draaien	ˈtaj̣ə	dʲajə

1;09.22	10	draaien	ˈtɛ̯ijə	təaːjə
1;09.22	10	brood	ˈʋoŭt	
1;09.22	10	brood	ʔəˈboːt	
1;10.03	11	broek	ˈbut	bəut
1;10.03	11	brood	ˈḅoːt	bəoːt
1;11.08	12	draaimolen	ˈtaimˈboˑjə	dɛaːimboːle
1;11.08	12	zebra	ˈteˑp	

- Broad transcriptions are identical to the CLPF transcription; narrow transcriptions are provided by the author of this thesis and focus on the representation of the cluster with greater phonetic detail.

Chapter 4. Two-year-olds' cluster productions in naming tasks

4.1. Introduction

In Chapter 2 the realization of target words starting with /Cr/ and/ kn/ clusters was studied, and it turned out that an acoustic trace of the omitted segment was present. Chapter 3 focused on the longitudinal spontaneous realization of target words starting with a /Cr/ cluster and a developmental pattern in the realization of these words was found, where the presence of an acoustic trace occurred in a specific developmental stage, preceded by a stage where no acoustic trace was present.

Up until now the data that were analyzed were mostly spontaneous utterances. In this chapter I report on a more experimental approach to longitudinal cluster production, the goal of which is to locate in a more controlled way the problematic levels of processing in the model and to get insight into the development of the production mechanism. In Chapter 1, the possible effects that malfunctioning/absent modules in the model of speech production may have on children's spontaneous word productions were discussed. Here I use the model to make predictions about the performance on different types of production tasks. The idea is that the performance on different types of production tasks, namely picture naming, word repetition and nonword repetition, can tell us something about the functioning of the different modules in the production mechanism. In a similar way, Den Ouden (2002) compared the performance of aphasic patients on production tasks. There too, the ultimate goal was to detect the layer in the speech production mechanism of each patient at which problems occurred that caused phonological errors. Since Den Ouden's study is one of the small number of studies in which the Levelt et al. (1999) production model is used to study a speech system deviating from the norm, and since child language data also show deviations from the norm, a similar study with two-year-olds was planned.

For the present study, the tasks used by Den Ouden were adapted to become suitable for two-year-old children. In addition, the production tasks were administered several times over a longer period of time in order to see whether changes occurred that could point to developmental changes in the production mechanism. The intention was to also include a longitudinal perception task, which would be able to tell us about the individual development of the lexical representation of the onset clusters of target words. However, due to problems with the design of the study, it turned out to be impossible to interpret the results of these experiments in a meaningful way. Unfortunately one source of information is therefore missing. The remainder of this chapter is organized as follows: In 4.2 I discuss the theoretical background of the present work and explain what performance on the different tasks can tell us about the developmental state of the production mechanism. In 4.3 the materials and methods of the different tasks are presented. In 4.4 the results of the individual children will be discussed in detail. A general discussion and conclusions are presented in 4.5.

4.2. Background

According to Kohn and Goodglass (1985), phonological errors of patients with aphasia could be the result of damage that causes problems either with lexical access, or with access to the functioning of phonological encoding, phonetic encoding or articulation. Following up on this idea, Den Ouden (2002) designed an experiment that aimed to trace the source of the segmental problems of patients with aphasia to lexical access, phonological encoding or phonetic encoding. He did not focus on the level of articulation because when problems occur at this level, it results in a particular kind of aphasic disorder, namely dysarthria of speech. Den Ouden designed three tasks, picture naming (PN), word repetition (WR) and a phoneme detection task (PERC), and explained in what way the scores on these tasks could be used to identify the functional locus of the impairment in the Levelt et al. (1999) speech production model. According to Den Ouden, deficits at a particular level result in a specific performance pattern in these tasks: if the impairment lies at the level of lexical access, patients will perform better on word repetition than on phoneme detection and picture naming, while performance will be poor on all three tasks if the functional locus of the impairment is at the phonological encoding level. Impairment at the level of phonetic encoding causes poor performance on the picture naming and repetition tasks, while phoneme detection should not be affected. This will be discussed in more detail below (4.1.2). I now first turn to some production studies with young children that have been performed previously, and are relevant to the present study.

4.2.1. Young children's performance on production tasks

In the literature, extensive attention has been paid to how children in different age groups perform on production tasks. Numerous acquisition studies have focused on the differences between naming and repetition tasks (Hoff et al., 2008; Zamuner, 2009; Munson et al., 2005), or differences between nonword repetition (NWR) and other measures of productive vocabulary (Metsala, 1999; Bowey 2001; Paradis, 2011). However, the focus of these studies was different from the focus of the present study, and either lay on the relation between phonological memory, as represented by the performance on a NWR task, and vocabulary size, or on the relation between phonotactic probability and production success. The most relevant studies for this chapter are the ones by Vance et al. (2005), Hoff et al. (2008) and Zamuner (2009).

The main goal of the study of Vance et al. (2005) was to test the speech production model by Stackhouse and Wells (1997), a model very similar to that of Levelt et al. (1999). In order to find out which part of the model is affected when children of different age groups make speech errors, PN, NWR and WR tasks are carried out with English-speaking children between 3 and 7 years of age, and for each age-group their performance on the three tasks was compared. Their responses were scored as being either correct or incorrect. For the 3-year-olds performed worse on the PN task than on the two repetition tasks, while the 4-year-olds performed worse on the PN and the NWR tasks as compared to their performance on the WR task. Not surprisingly, the older the children were, the better their performance on the PN task became. The authors interpreted the poor performance on the PN task by 3-year-olds, as resulting from problems retrieving the words from the mental lexicon. They performed better on the repetition tasks because they were aided by the presence of the adult model. In the 4-year-olds, some immaturity of the lexical representation still affected the performance on the PN task, which was worse than their performance on the WR task. In the performance of the 5-year-olds, the difference between WR and PN had disappeared, while they continued being less accurate on the NWR task, just like the 6- and 7-year-olds. The authors suggest that for the oldest age groups there is a beneficial effect of the lexical representation on speech output processing. It appears that the speech processing requirements of discriminating all the phonemes of the nonword, without top-down support of the mental lexicon, and with the additional task of creating a new motor program, negatively affect the performance on the NWR task.

Since the children studied in this thesis are around two-years old, the study by Hoff et al. (2008) is relevant. Here, two groups of English-speaking children, 20- and 24-month-olds, were tested. These children's real word and nonword repetitions were assessed, together with their productive vocabulary. The PCC (percent consonant correct) was calculated for the children's productions. According to this measure, the percentage of correct consonants in a word is calculated (number of correct consonants / total number of consonants × 100, where a consonant that has been substituted or deleted obtains zero points, while a correct consonant obtains one point). The vocabulary size was measured with the MacArthur-Bates Communicative Development Inventory CDI.

The results in Hoff et al. (2008) showed that the 20-month-olds scored significantly worse on the NWR than on the WR task, and that performance on the NWR task and vocabulary size were strongly correlated. These results were replicated with the 24-month-olds. The authors of this study conclude that NWR-accuracy reflects phonological memory capacity and that this capacity is related to the level of vocabulary development of children.

In the study by Hoff et al. (2008) the nonwords were phonologically matched to the real words but they were not controlled for their phonotactic probability. Zamuner (2009) tested the production of nonwords of 28 and 31-month-old Dutch speaking children. The stimuli consisted of nonwords that varied in the degree of phonotactic probability (PP) of the consonants in onset or coda position. The nonwords either had an onset or a coda with a low phonotactic probability, or an onset or a coda with a high phonotactic probability. Zamuner controlled for the neighborhood density of the constructed stimuli and found out that there were more neighbors for the high probability nonwords and more neighbors for nonwords differing in segments in word-initial position. The responses were scored as correct, incorrect or as no response. The analyses were based on the proportion correct responses per nonword category (low PP onset, low PP coda, high PP onset, high PP coda).

The first main finding was that phonotactic probabilities influenced children's accuracy in the production of nonwords, both in word onsets and in word codas. Children produced nonwords with high phonotactic probability more accurately, independent of the position. The second finding of importance was that children's vocabulary size correlated with the accuracy of their production. More specifically, children with larger vocabularies were more accurate in the production of segments in word onset position. This effect was explained by the higher neighborhood density for lexical items contrasting in word onset position. If more lexical items contrast in word initial position, then phonological representations of this position should be more developed, according to Zamuner.

From these studies we can conclude that children as young as 20 months are able to perform on PN and (N)WR tasks. For this young age-group, performance on these tasks has up until now only been correlated with vocabulary size and phonological memory, but not with the developmental state of the speech production mechanism. We will now turn to this mechanism again, and discuss, along the lines of Den Ouden (2002), the expected performance on production and perception tasks of two-year olds, given the potential developmental problems with lexical access, phonological encoding or phonetic encoding.

4.2.2. The (developmental) state of the production mechanism and performance on different tasks

4.2.2.1. The level of lexical access

The mental lexicon of a two-year-old child is still under construction and it is likely that stored forms are not always completely or correctly specified. Evidence from experimental infant perception studies sometimes points to detailed phonetic representations, and sometimes to incomplete phonetic specifications, depending on the age of the infants and the position of the segment in the word (Fikkert, 1994; Levelt, 2012; Stager & Werker, 1997; Swingley, 2009; Trehub et al., 2007; Zamuner, 2009;). As discussed in Chapter 1, an incorrect representation is expected to lead to regular incorrect word productions, while an underspecified representation could lead to variable word productions. A child who has problems at this level is expected to have problems with the PN task. In a naming task, the speaker needs to consult his or her mental lexicon in order to find the stored form that goes with the depicted object. In case an incorrect form is stored, an incorrect form will be produced. In a repetition task the lexical representation is not necessarily activated, since the auditory form is provided. Performance on a WR task could thus be better than performance on a PN task when problems lie at the level of lexical access or the stored lexical representation. It is possible, however, that the child does activate the lexical representation of a known word during repetition, blurring the difference between the two tasks. However, this route seems to be blocked in the nonword repetition task, since in the case of nonword repetition, there is no existing word form stored in the mental lexicon. Although it has been shown that even nonwords can activate the lexicon through word-likeness (Swingley & Aslin, 2000; Zamuner, 2009), performance on this task is expected to be largely unaffected when the level of lexical access is the source of the deviating word productions. Finally, if the lexical storage is incorrect, or if lexical access is problematic for a child, it should be difficult to perceive subtle differences between words - like between the correct form [trɛin] for *trein* (train) and the simplified form [tɛin]. In other words, we expect poor performance on a young children's version of Den Ouden's phoneme perception task.[1] To summarize, good performance on the NWR task in combination with poor performance on the PN (and PERC) task(s) would point to problems at the level of lexical access. Performance on the WR task could either be comparable to performance on the NWR task or to performance on the PN task, depending on whether the lexical representation of the to-be-repeated form is activated are not. In short:

Lexical Access/Representation Problem:
NWR, WR >> PN, PERC
or
NWR >> PN, WR, PERC

The conclusion reached by Vance et al. (2005) for the performance of the 3- and 4-year-olds, namely that the better performance on repetition tasks than on the

[1] Note that this task is not meant to test a child's general auditory perception abilities, but his/her linguistic perception abilities.

PN task entails a lexical retrieval problem, thus very closely resembles[2] the above reasoning of Den Ouden (2002) for a potential source of problems.

4.2.2.2. The level of phonological encoding

At the level of phonological encoding, the sounds of the activated lexical item are retrieved and syllabified. At this level, then, an underlying *segmental* representation is mapped onto a *phonological output* representation. In picture naming, after retrieving the lexical item from the lexicon, this item needs to pass through the phonological encoding module in order to be produced. In the case of repetition, the phonological encoding stage can either be skipped, when the lexicon is bypassed, or not, in case the lexical route is taken.

Problems at the level of phonological encoding are not expected to affect the performance on a perception task (Den Ouden, 2002). If a child has stored a target-like segmental representation in his or her mental lexicon, he or she should be able to perform well on a perception task, despite a deficit at the phonological encoding level.

To summarize, poor performance on the PN task(s) in combination with good performance on PERC tasks is expected when there are problems at the level of phonological encoding. NWR could be good, when phonological encoding is bypassed, and WR could again either go with PN (poor) or with NWR (good). In short:

Phonological encoding problem:
NWR, PERC >> PN, WR
or
NWR, WR, PERC >> PN

---

[2] There is no reference to Den Ouden (2002) in Vance et al. (2005).

In order to differentiate a lexical access problem from a phonological encoding problem, performance on the PERC task is crucial. If PERC goes with NWR, the problem source is phonological encoding, while if NWR is better than PERC, then the problem source is lexical access. Since there is no PERC test in Vance et al. (2005) to differentiate the two sources, their 3- and 4-year-olds could also have had problems at the phonological encoding level. Unfortunately, because of the case study nature of the experiment, the PERC task I used could not give meaningful results and was left out. Therefore I only collected meaningful data from the children's performance on the production tasks.

4.2.2.3. The level of phonetic encoding

During phonetic encoding, a motor program is constructed and the phonemic string is mapped to gestural commands. This also requires the awareness of language-specific allophonic details of each sound. When a string of sounds is repeated, the acoustic form is directly translated into a gestural score at this level (Browman & Goldstein, 1989; Boersma, 1998). If there are problems at the level of phonetic encoding, all production tasks will be affected. The PERC task will remain unaffected, for the same reasons as given above for the phonological encoding level. In short:

Phonetic encoding problem:
PERC >> PN, WR, NWR

4.2.2.4. The level of motor programing

Den Ouden does not discuss what the consequences for the model would be when we would find better performance on the PN task compared to performance on the WR task. Nijland and Maasen (2005) distinguish between imitation and spontaneous speech, where imitation is a synonym for both WR and NWR and spontaneous speech is a synonym for PN. They discuss the possible scenario that children might be able to produce known words in spontaneous speech while being unable to imitate them. According to the authors, this could arise, due to the fact that in spontaneous speech uttered words are “overlearned”, while during imitation, on-line contextual adaptation of the segments is required. Nijland & Maasen label this as a problem of motor programing since it specifically concerns the articulatory cohesion within a syllable. If the lexical route is taken in the WR task, then we would expect both PN and WR to outperform NWR. This resembles the situation of the 5-year-olds in the Vance et al. (2005) study. In short:

Motor programing problem:
PN >> WR, NWR
or
PN, WR >> NWR

To conclude this section, in a similar way as in Den Ouden (2002) I have described the different repercussions for the performance on PN, WR and NWR tasks, when a deficit at one of the three modules - lexical access, phonological encoding and phonetic encoding – is assumed.

4.3. Materials and methods

4.3.1. Participants

Six children participated in the longitudinal study, four girls and two boys. The data of two of the girls were not included in the study because one girl was bilingual and another girl consistently refused to participate in the nonword repetition task. The data presented here thus come from four monolingual Dutch children, two boys, Lars and Matteo, and two girls, Meike and Hannah. They completed all tasks in all sessions, but due to technical issues the recordings of Meike’s session 3 and Matteo’s session 2 were not stored properly and were therefore lost. Lars was recorded between the age of 1;7 and 2;7; Matteo was recorded between age 2;00 and 2;5; Meike was recorded between age 1;11 and 2;3 and Hannah was recorded between age 2;1 and 2;6.

Data collection for a child was terminated when at least one of the cluster types in which we were interested, /Cr/, /Cl/ or /sC/, was acquired. The recordings were carried out in the children's homes, usually in the living room, which was maintained as quiet as possible. All recordings were performed by myself.

4.3.2. Procedure

Each child was recorded in his or her home for at least five consecutive sessions. The children's utterances were recorded with a Microtrack II digital recorder and an external Microtrack II microphone. Each session was carried out as follows: first the PN task was conducted, using a powerpoint slide show on a laptop, followed by the WR task and the NWR (or viceversa), during which the laptop was closed.

4.3.3. Material

The words used in the PN and in the WR tasks were identical. The words used in the NWR task were based on the phonological form of the words in the real word tasks. See Table 1 for the list of words and nonwords used in the three production tasks. The stimuli were subdivided into stimuli containing the following cluster types: /Cr/; /fric+r/; /sC/; /s+fric/; /Cl/; /fric+l/; /tʋ/ and /kn/, where C in this chapter is used for a plosive.

In Figure 1 is an example of one of the pictures I used in the PN task. For the WR task, I produced the Dutch word myself and tried to elicit repetition by using the following phrases:

1. Zeg maar *trein*. (Say *train*.)
2. Kun jij *trein* zeggen? (Can you say *train*?)

Figure 1: A picture of a Dutch train, familiar to two-year-olds, used in the picture naming task.

Figure 2: Two objects which were new and therefore unknown to young children used (when necessary) in the nonword repetition task, which represent two microbes (giardia and e-coli), the size of a small teddy bear.

1: Words and nonwords used in the three production tasks (PN, WR and NWR tasks); Dutch orthography is used for the annotation.

Custer Types	Clusters	Words	Translation	Nonwords[3]
/Cr/	/dʀ/	draakje	dragon	droon
	/kʀ/	kraan/ kroon	faucet/ crown	kriep/ kraak
	/bʀ/	broek	trousers	braak
	/tʀ/	trein	train	traak
/fric+r/	/χʀ/	gras	grass	graak
	/fʀ/	fruit	fruit	friep
/sC/	/sp/	speeltuin	playground	spaam
	/sk/	skippybal	skippyball	skaam
/s+fric/	/sχ/	schaap/ schaar/ schoen	sheep/ scissors/ shoe	schaag
	/sʋ/	zwembad/ zwart	swimming pool/ black	zwiep
/sn/	/sn/	snoep	candy	snaak
/Cl/	/kl/	klok	clock	klot
	/bl/	bloem	flour	bliep
/fric+l/	/fl/	vlinder	butterfly	vloon
	/fl/	fles	bottle	flaak
	/χl/	glas	glass (cup)	gler
/tʋ/ /kn/	/tʋ/	twee	two	twot
	/kn/	knoop	button	knaak

[3] Some of the nonwords are low frequency, often old-fashioned real words that are unknown to the children in this sample.

For the NWR task, the child was first simply asked to repeat a specific nonword. However, in case this did not elicit any production from the child, an unknown object was shown to him or her (see Figure 2). This object was given a name, the nonword, and the child was asked to repeat this name (Hoff et al., 2008). The following elicitation phrase was used in this case:

> Kijk, dit is een *traas* Hoe heet hij (ook al weer)?
> (Look this is a *traas*, what is its name (again)?)

The list of real words used in the PN, the WR tasks and the nonwords used in the NWR task are presented in Table 1.

For this test I was specifically interested in the effect of the different production tasks on the children's performance on cluster production. The words and the non-words that were compared therefore had to have similar phonotactic probabilities. To this end I computed the diphone transitional probabilities of the words and the nonwords based on the CELEX corpus of the Dutch language. After computing the diphone transitional probabilities, an averaged log transitional probability was obtained (Adriaans, 2011). Words for the WR and the PN tasks were considered suitable stimuli when they fulfilled three requirements. First, the selected words had to be familiar to two-year-olds, secondly, they had to be easy to visualize and, finally, the words had to start with different types of onset clusters. The different requirements made it difficult to keep the transitional probabilities (TPs) identical for all real-word/non-word pairs of stimuli. In Appendix 5[4], the TPs of the 22 real words and the 19 nonwords are presented. The mean log TP of the real words is -1.23, ranging between -1.49 and -0.90. The word with the highest logarithmic transitional probability in our list of words (-0.90), is *vlinder* (butterfly); while

[4] In the cluster types /sʋ/ and /sχ/, the TP only of the first word in the list reported in Table 1 was taken into consideration.

the word with the lowest log TP is *snoep* (candy). The real words are part of the first 1000 words from the obligatory vocabulary for Dutch preschool children (Bacchini et al., 2005).

The low frequency words go together with a high log TP, while the high frequency words go together with low log TP. For instance the word *vlinder*, is low frequent and meanwhile is also characterized by higher log TP (taking into account its negativity). The word *snoep*, on the other hand has a high frequency and a low log TP.

For the nonwords in our stimuli set, the mean log TP was -1.22, ranging between -1.42 and -1.11, where the high log TP of the word *zwiep* is an indication that, if it were a word, *zwiep* would be a word of a low frequency, while *braak,* with its low log TP of -1.42 would be a highly frequent word. We carried out a paired sample t-test to compare the log TPs of the real words with those of the nonwords and found no significant difference between the two sets of words ($p > .1$).

4.4. Results

4.4.1. Quantitative analysis

The children's responses were first phonetically transcribed by an experienced transcriber and subsequently they were categorized either as containing a complex onset cluster or not. Since I was especially interested in the acquisition of onset clusters, the accuracy of the segments following the onset cluster was not scored. I therefore did not use measures like PCC, *percent consonant correct*, (Shriberg & Kwiatkowski, 1982) and the PCC-R, *percent consonant correct – revised*, (Shriberg et al. 1993; Shriberg, et al. 1997), where both deletions and substitutions are scored as errors. Here I consider a cluster to be acquired when a sequence of two consonants is realized. This means that a cluster produced with consonant substitution (disregarding whether the substituted consonant is C1, C2 or both) also counts as a cluster that has been acquired[5]. In the general analysis presented here, I make a distinction between cluster omission and cluster reduction. Below in Figure 3 I present the percentages of reduced (CV) and complex (CCV) clusters per session, per task, per child.

The performance of the four children over time on the PN task is shown in Figure 3, where a graph of the percentages of reduced [CV] and complex [CCV] realizations of the target onset clusters are presented for all four children.[6] In general, the same picture emerges for the other tasks, with a slightly different timing.

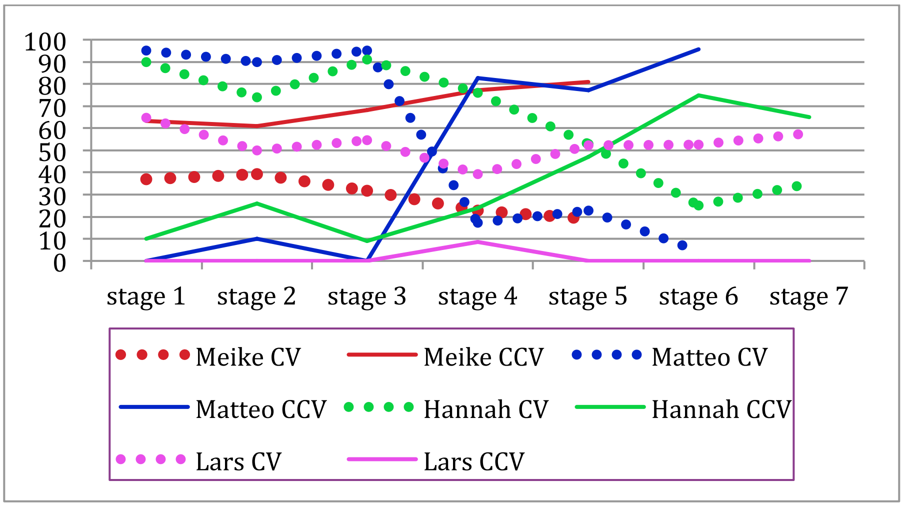


Figure 3: Percentage of the cluster realizations as /CV/ utterances (dotted line) and /CCV/ utterances (straight line) in the Picture Naming task by the four children.

[6] See Appendixes I to IV for all transcriptions of the data of all four children.

The general picture that can be deduced from Figure 3 is, as expected, that the number of reduced [CV] realizations decreases over time, while the number of complex [CCV] realizations increases. For two children, Matteo and Hannah, there is a clear breakpoint – at stage (here session) 4 for Matteo, and at stage (session) 6 for Hannah – while for Meike this breakpoint seems to have occurred already at some point before the first recording session. Lars, finally, is not really making progress in his realizations of complex [CCV] in the PN task at all in the data collecting period.

Graphs based on the percentages of realized CCV utterances of the individual children in the different tasks are presented in Figures 4-8. Since Lars hardly showed any development from reduced CV to complex CCV, but did show a development from omitted ØV to reduced CV, in his graph below the CV realizations are depicted. Here we see that the children perform differently in the different tasks, and that initially the highest percentages of [CCV] (or CV for Lars) realizations are found in the NWR task. In the final recordings, performance on the different tasks is more or less equal. For Matteo, performance on the WR task is similar to the performance on the PN task, while for Hannah, in the course of development, performance on the WR task becomes similar to the performance on the NWR task. For Meike performance on PN and NWR shows a similar pattern, while the performance on WR lags behind for some time.

Lars exhibits low percentage of ØV but a high percentage of CV forms in the first sessions in the NWR task. Overall, the word tasks show poorer performance in the first sessions (more ØV forms) and better performance in the final sessions (more CV forms). In the final session all tasks show an occurrence of ØV forms of around 45% and an occurrence of CV forms of around 55%.

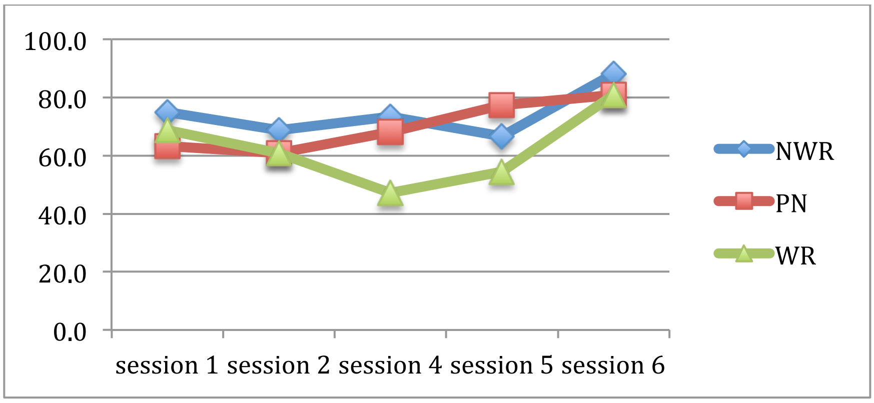


Figure 4: Percentage /CCV/ realizations in the NWR, PN, WR tasks for Meike.

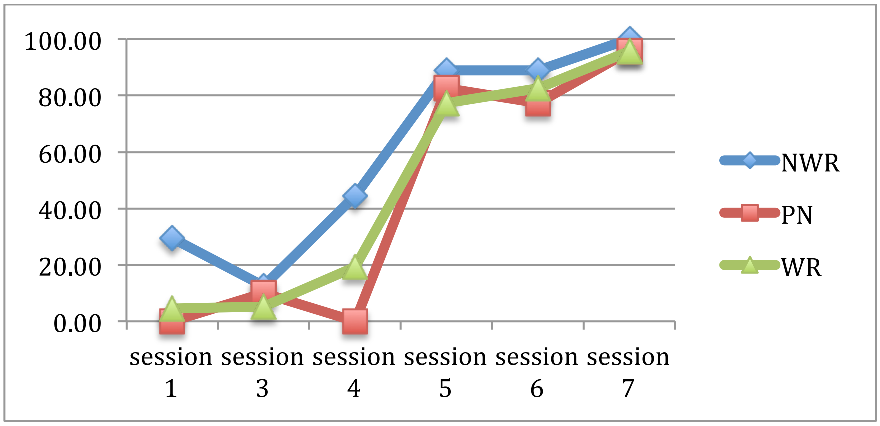


Figure 5: Percentage /CCV/ realizations in the NWR, PN, WR tasks for Matteo.

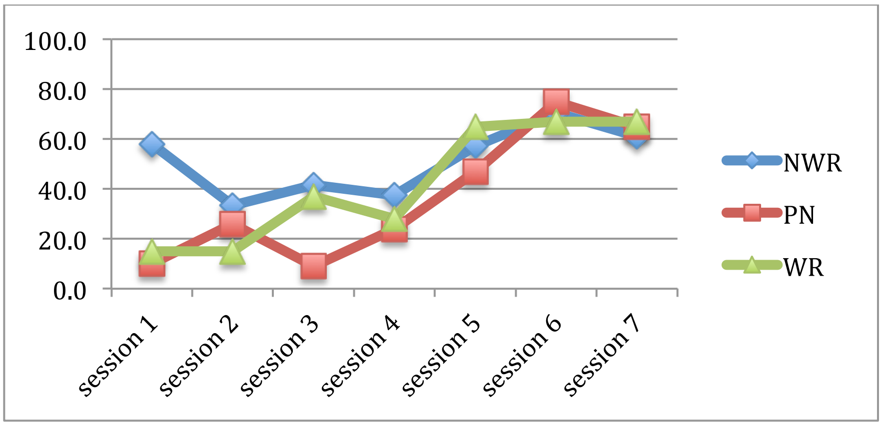


Figure 6: Percentage /CCV/ realizations in the NWR, PN, WR tasks for Hannah.

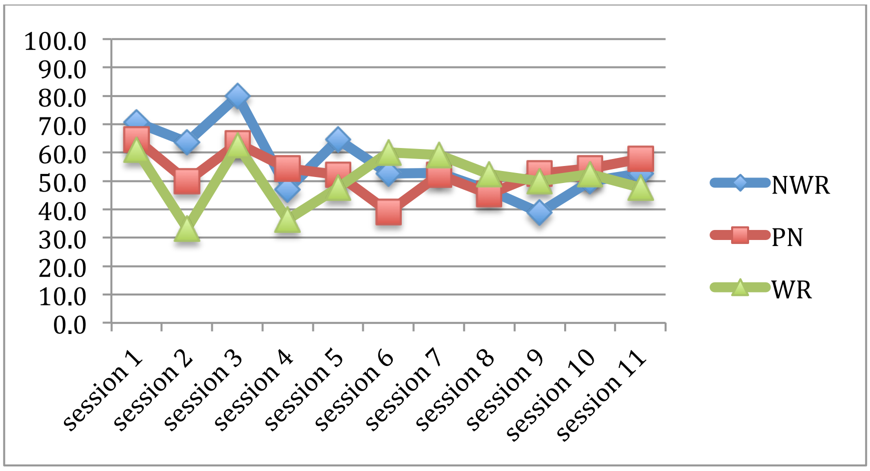


Figure 7: Percentage /CV/ realizations in the NWR, PN, WR tasks for Lars.

4.4.2. Intermediate summary

Two general patterns emerge from the data. The first salient pattern is that initially the highest percentages of cluster realizations (or singleton consonant realizations for Lars, see below) are found in the NWR task. The second general pattern is that in the final recordings, performance on the three different tasks is very similar. Except for Lars, performance on all tasks also shows a steady (Hannah, Matteo) or a more gradual improvement (Meike).

As explained in 4.2.2.1, if children score better on the NWR task than on the word task, then we can conclude that they have problems either with lexical access or with the lexical representation itself. In the case of NWR, the nonwords lack a representation in the mental lexicon, and this is why neither an incomplete lexical representation nor a phonological encoding problem could negatively affect the production of NWR items. Only real word productions, and real word repetitions in case the lexical route is taken, can be negatively affected.

Another finding is that some children appear to take the lexical route in the WR task, and therefore show similar performance on the WR and PN tasks (Matteo), while others (Hannah, Lars) appear to take the non-lexical route in the WR task, and perform in a similar way on the NWR and WR tasks. For Meike neither route can explain her results, since performance on the PN and NWR tasks is similar, while WR exhibits the poorest performance. In the discussion I will try to come up with an explanation for her poor performance on the WR task.

I will now turn to the results of the individual children, and discuss their performance on the different tasks and development in more detail.

4.4.3. Qualitative analysis

In the paragraphs to come I will offer an explorative analysis of the linguistic and psycholinguistic patterns found in the speech development of each individual child. The relatively small amounts of data within each session, within each production task and for each child preclude a statistical analysis. However, the results from our exploratory analysis do give an additional preliminary insight into the development of the speech production mechanism, and can be used to set up future research.

4.4.3.1. Case study Meike (1;11 - 2;3)

For Meike, the data of 5 recordings could be analyzed. She produces reduced versions of the cluster types /sC/, /kn/ and /zʊ/ in the first session, and still reduces /sk/, /sp/, /sx/ /zʊ/ in the final session. Production of the clusters /sl/, /sn/, /kn/ and /tʊ/ shows development over the sessions. For all Meike's productions see Appendix 1.

In Table 2 are the number of cluster realizations per session (raw numbers), the total number of productions (in parentheses) and the percentage of cluster realizations in the NWR, WR and PN tasks.

Table 2: Cluster realizations by Meike in the different tasks

	Sess1	Sess2	Sess4	Sess5	Sess6
NWR	12 (16) 75%	11 (16) 68.8%	11 (15) 73.3%	12 (18) 66.7%	15 (17) 88.2%
WR	11 (16) 68.8%	14 (23) 60.9%	8 (17) 47.1%	12 (22) 54.5%	17 (21) 81%
PN	12 (19) 63.2%	14 (23) 60.9%	15 (22) 68.2%	17 (22) 77.3%	17 (21) 81%

Three developmental stages can be discerned: a first stage formed by sessions 1 and 2, a second stage formed by sessions 4 and 5 and a third and final stage in the last session. In the first stage, the performance on the NWR task is better than on the two real word tasks (PN and WR). In the second stage, both PN and NWR show higher cluster realization scores than WR. Finally, in session 6, performance on all three tasks is similar, and the percentage of target-like cluster realizations is high, above 80%.

Compared to the general pattern described above in 4.4.2, the main difference is that the low scoring on the WR task compared to the other tasks in the second stage (sessions 4 and 5) makes it impossible to categorize Meike as either a lexical route-taker or a non-lexical route-taker in the WR task. However, if we look at the actual forms that are uttered in the WR, PN and NWR tasks in session 4, there are hardly any target clusters that are produced correctly in the PN or NWR task, but are reduced in the WR task – there is only one case where Meike performs better in both the NWR and the PN task (NWR *knaak* [kna:k], PN *knoopjes* [klo:pjəs], WR *knoopjes* [no:pjəs]) and two cases where PN is better than WR (PN *twee* [dve], WR *twee* [ve:] and PN *kroon* [kro:n], WR *kroon* [xo:n]). The words *knoopjes* and *kroon* were produced with a correct cluster in the previous – and following – sessions in the WR task, while the cluster in PN *twee* was reduced in the previous and following sessions. The apparent discrepancy between NWR and WR, or PN and WR in session 4 is thus not so obvious when we look at the actual productions. This is very different from the discrepancies between conditions in the other children's data. For example, in session 4 Matteo utters no forms with clusters at all in the PN task, compared to eight cluster productions in the NWR task. In Meike's session 5, however, there are four cases where performance on the PN task is better than on the WR task, all involving the sound /x/ - in /sx/ or /xr/ clusters. This could mean that Meike does not take the lexical route in the WR task, and that cluster production in the PN task is facilitated by the activation of the segmental representation of the word.

In general, Meike produces stable and segmentally correct clusters from the start for most of the Cr/Cl clusters in all tasks. All /sC/ clusters are problematic for Meike. Since /sC/ clusters violate the sonority sequencing principle for onsets when C is an obstruent – consonant sequences in the onset should have increasing sonority – it has been proposed that /s/ in these clusters occupies an "extra-syllabic position" (ESP, Kager & Zonneveld 1986). Obstruent-liquid clusters and /s/ + obstruent clusters thus have different syllabic representations. Fikkert (1994) has shown that children vary in the order in which they acquire these different cluster types: some children acquire obstruent-liquid clusters first, while others acquire the /s/+ obstruent clusters first. In principle, /s/+ sonorant clusters could receive either a complex onset representation, since they obey the sonority sequencing principle, or they could be grouped with the /s/ + obstruent clusters and receive an ESP representation. Children seem to vary in the way they group these /s/ + sonorant sequences, and they either acquire these sequences simultaneously with other fricative + sonorant clusters, or simultaneously with /s/ + obstruent clusters (Fikkert 1994). The fact that Meike has problems with all /sC/ clusters, while other fricative + liquid clusters are produced correctly shows that she groups /s/ + sonorant clusters with the /s/ + obstruent clusters. Syllabification takes place at the level of phonological encoding. It can thus be expected that as long as the "extra-syllabic-position" is not acquired, or not available, the /s/ cannot be syllabified, and will not receive a motor program. As a result the /s/ will not be produced. This would affect the production of /sC/ clusters in the PN task, but not necessarily in the repetition tasks. The first (correct) cluster productions of target /sC/-cluster words do indeed appear in the NWR and WR tasks. As soon as the ESP representation is available for phonological encoding of a sequence of consonants, this is expected to facilitate the production of /sC/ clusters in the PN task, but again the repetition tasks will not necessarily be positively affected; performance could now even be worse in the repetition tasks than in the PN task. This is what we appear to see with Meike's production of /sx/ clusters in session 5, described above. Performance on the NWR and WR tasks – if the non-lexical route is taken – thus seems to be unstable, unlike performance on the PN task. In this task, productions will systematically go wrong when the representation is incomplete or when phonological encoding is problematic, but there will be systematic improvement when developments have taken place at these levels.

4.4.3.2. Case study Matteo (2;0 - 2;5)

Matteo was recorded between the age of 2;0 and the age of 2;5, and 6 out of 7 recording sessions could be analyzed. Matteo produced reduced productions of all tested cluster types in the initial session, and had acquired all of them by the time of the final session.

In Table 3 are the number of cluster realizations per session (raw numbers), the total number of productions (in parentheses) and the percentage of cluster realizations in the NWR, WR and PN tasks, for Matteo.

Table 3: Cluster realizations by Matteo in the different tasks

	Sess1	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	5(17) 29.4%	1(8) 12.5%	8(18) 44.4%	16(18) 88.9%	16(18) 88.9%	19(19) 100%
WR	1(22) 4.5%	1(19) 5.3%	5(21) 23.8%	17(22) 77.3%	19(23) 82.6%	22(23) 95.7%
PN	0(21) 0%	2(20) 10%	0(21) 0%	19(23) 82.6%	17(22) 77.3%	22(23) 95.7%

There appear to be three developmental stages, formed by sessions 1-4, 5-6, and 7. In sessions 1-4 the performance on both the PN and WR tasks is very low, in sessions 5-6 there is a break-through and performance is suddenly high on all tasks, and in session 7 performance is almost at ceiling. Throughout the sessions, the number of cluster realizations is remarkably high in the NWR task (with exception of session 3). In 9 out of 19 cases where items are produced in all three tasks, the first cluster production occurred in the NWR task – in 9 cases the cluster appeared in all three tasks in the same session and in 1 case (*kraan*) a cluster production appeared in the WR task first. The largest difference between PN and NWR is in session 4. Performance on WR goes with the performance on PN, which suggests that Matteo takes the lexical route in the WR task.

For Meike a clear difference in development between /sC/ clusters and other clusters was found. This is less clear in Matteo's case, where all clusters seem to show up in the PN task at the same time, in session 5. However, target /sC/ clusters are the first to receive – usually incorrect – cluster productions in the NWR task. As mentioned above, it is actually not expected that the different phonological representations, ESP position versus complex onset, will play a role in repetition tasks like NWR. For Matteo, then, the initial /s/ could have acoustically highlighted the fact that a sequence of consonants should be produced. The fact that target /sp/ is the first cluster to be produced in a stable and correct way in the PN task, from session 3 on, could mean that this sensitivity, in turn, caused the early development of ESP processing during phonological encoding for Matteo. I will come back to this in the discussion.

4.4.3.3. Case study Hannah (2;1-2;6)

Hannah was recorded for 7 sessions between the age of 2;1 and 2;6, and all sessions could be analyzed. Except for the target clusters /xl/ and /sl/, she reduced all cluster types in the first recording session, and still reduced almost all /Cr/ clusters in the final session.

In Table 4 are the number of cluster realizations per session (raw numbers), the total number of productions (in parentheses) and the percentage of cluster realizations in the NWR, WR and PN tasks, for Hannah.

Table 4: Cluster realizations by Hannah in the different tasks

	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	11(19) 57.9%	5(15) 33.3%	5(12) 41.7%	6(16) 37.5%	8(14) 57.1%	12(17) 70.6%	11(18) 61.1%
WR	3(20) 15%	3(20) 15%	5(13) 37%	5(18) 28%	13(20) 60%	14(21) 67%	14(21) 67%
PN	2(21) 10%	5(19) 26%	2(22) 9%	4(17) 24%	9(19) 47%	15(23) 65%	15(23) 65%

Three developmental stages can be discerned, one formed by sessions 1-2, one by session 3-5 and one by sessions 6+7. The first 4 sessions are similar in the sense that performance is best on the NWR task and worst in the PN. In 11 out of the 16 cases where targets are produced in all three tasks, the first cluster production, correct or incorrect, occurred in the NWR task. In 5 out of 11 cases where the cluster was eventually produced correctly this occurred in the NWR task first, in 5 cases the correct cluster occurred in all tasks at once, and in 1 case it occurred in the WR task first. Thus, like in the cases of Meike and Matteo, the NWR task exhibits the most adult-like cluster realizations. In session 3, performance on the WR task becomes much better, and performance on WR and NWR outranks performance on PN. In the final two sessions performance is similar in all three tasks.

Cluster productions of target /s/+ obstruent clusters appear somewhat later than cluster productions of C+liquid clusters. All target /s/ + obstruent clusters are produced [st] in the final session. Target /sl/ goes with the other /Cl/ target clusters and appears (correctly produced) early. Target /sn/ goes with the target /s/+ obstruent clusters, but is produced correctly in the PN task from session 5 on. It thus appears that the ESP is acquired somewhat later than the complex onset, affecting the PN and WR tasks, but not the NWR task, where early – incorrect – clusters are produced for target /sC/ clusters.

4.4.3.4. Case study Lars (1;8-2;7)

For Lars 11 recording sessions are available, between the age of 1;8 and the age of 2;7, and all sessions could be analyzed. The data from all sessions are presented in Appendix 4. Lars has a different developmental pattern from the other children. While he reduced target clusters starting with a plosive, and /s/ + plosive clusters to singleton plosive consonants, he omitted the entire cluster from his productions if the target cluster was a fricative + liquid cluster. The complete omission of these target clusters is related to the fact that he also omitted target singleton fricatives and target singleton liquids from his productions (zeep /zep/ *soap* becomes [ep], goed /xut/ *good* becomes [ut], rood /rot/ *red* becomes [ot] for example). By the end of the recording period he still does not produce any consonant cluster spontaneously, and there are only a handful of instances where he produces a cluster in a repetition task.[7] Table 5 shows the number of cluster realizations, while Tables 6 and 7 show the number of singleton consonant realizations for target clusters starting with a plosive or with sP (Table 6), and for target clusters starting with a fricative and other s-clusters (Table 7), again per session (raw numbers), the total number of productions (in parentheses) and the percentage in the NWR, WR and PN tasks. Target cluster words that are not realized with a singleton consonant are, with only a few exceptions that are indicated, produced with completely omitted clusters.

[7] Lars started producing all consonant clusters in kindergarten and is fine now.

Table 5: Number of /CCV/ realizations per session

	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6
NWR	0(17) 0%	1(11) 9.1%	0(5) 0%	0(17) 0%	0(17) 0%	1(19) 5.3%
WR	0(18) 0%	0(14) 0%	0(8) 0%	2(22) 9.1%	0(21) 0%	0(5) 0%
PN	0(17) 0%	0(3) 0%	1(19) 5.3%	0(22) 0%	0(23) 0%	2(23) 8.7%

	Sess7	Sess8	Sess9	Sess10	Sess11
NWR	2(17) 11.8%	0(15) 0%	0(18) 0%	1(18) 5.6%	0(19) 0%
WR	0(22) 0%	0(21) 0%	0(22) 0%	1(21) 4.8%	1(21) 4.8%
PN	0(23) 0%	0(22) 0%	0(19) 0%	1(22) 4.5%	0(19) 0%

Table 6: Number of /CV/ realizations for plosive-initial and sP target clusters

	Sess1	Sess2	Sess3	Sess4	Sess5
NWR	9(9) 100%	6(6) 100%	3(3) 100%	8(8) 100%	9(9) 100%
WR	10(10) 100%	6(6) 100%	9(9) 100%	10(10) 100%	10(10) 100%
PN	9(9) 100%	1(1) 100%	6(6) 100%	8(10) 80% *	10(10) 100%

	Sess7	Sess8	Sess9	Sess10	Sess11
NWR	7(8) 87.5%*	7(7) 100%	7(8) 87.5%	8(9) 88.8%*	9(10) 90%
WR	11(11) 100%	9(9) 100%	8(8) 100%	10(11) 90.9%*	10(10) 100%
PN	10(10) 100%	10(10) 100%	10(11) 9.9%	10(11) 90.9%	9(10) 90%

*there is 1 CCV realization in this session **there are 2 CCV realizations in this session

Table 7. Number of /CV/ realizations for fricative-initial clusters (except sP)

	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6
NWR	2(7) 28%	1(4) 25%*	1(2) 50%	0(9) 0%	2(9) 22.2%	1(10) 10%
WR	0(6) 0%	0(5) 0%	2(7) 28.5%*	0(8) 0%	0(9) 0%	1(9) 11.1%
PN	0(7) 0%	0(2) 0%	0(2) 0%	0(9) 0%	0(9) 0%	0(4) 0%*

	Sess7	Sess8	Sess9	Sess10	Sess11
NWR	2(9) 22,2%*	0(8) 0%	0(8) 0%	1(9) 11.1%	0(8) 0%
WR	0(8) 0%	0(8) 0%	0(9) 0%	1(7) 14.2%*	0(7) 0%
PN	1(9) 11.1%	0(8) 0%	1(9) 11.1%	0(7) 0%	1(18) 12.5%

*there is also 1 CCV realization in this session

Unlike the other children, Lars does not show any systematic development towards cluster production, as can be seen in the Tables 5-7. However, if anything, it is again clear that the more advanced productions show up in the repetition tasks, and predominantly in the NWR task. Because of the lack of development in general with respect to cluster production, in Lars' case we could not find any indication of a different development for sC and C+liquid clusters.

4.6. Discussion

Den Ouden (2002) tried to pinpoint the source of errors in aphasic speech. He did this by administering a series of tests to speakers with aphasia, Non-word Repetition, Word Repetition, Picture Naming and a perception test. The scores on the different tests together indicated where in the speech production model the error source was located for an individual speaker. The aim of the present study was to apply this method to pinpoint the source of errors in young children's speech. While the speech production mechanism of an aphasic speaker is – at some point – stable, the speech production mechanism of a child is under construction, and changes over time. Therefore, by following young speakers for a longer period of time, I hoped to find evidence for the way in which the speech production mechanism develops; a change in the relative ranking of performance on the different tasks can be taken to indicate maturation of one of the encoding modules, and in the end it is expected that performance is equally good on all tasks. The focus of the present study lay on onset clusters, since these are error-prone in young children, and show a gradual development.

With respect to performance on the different tasks over time, there is indeed a developmental shift (except in the case of Lars), from an initial state where performance on NWR outranks performance on both WR and PN, to a final state where performance on all tasks is similar, and has improved, in the final sessions. In terms of Den Ouden (2002), then, in the initial state the problem lies either with lexical access or with the lexical representation, or with phonological encoding. Due to the fact that perception data are lacking, I cannot distinguish between these two possibilities. The non-lexical route through the speech production system is initially more successful than the lexical route, and performance on the WR task depends on the route that is taken in this task. The intermediate stage of Meike, where PN outranks both repetition tasks, shows that as soon as problems with lexical access/the lexical representation or phonological encoding are solved, the lexical route can actually boost performance because it quickly leads the speaker to the 'learned' motor program that is associated with the representation. Repetition requires constructing a new motor program, which might lead to errors. This is why performance on repetition tasks can be more successful initially, when difficulties with lexical access, struggling with an underspecified representation, or poor execution at the level of phonological encoding can be avoided, but less successful at later stages.

While the information on the children's performance on a perception task, which could be used to differentiate between a lexical access problem or a phonological encoding problem, is missing, the different performance on target /sC/ clusters and target C+liq clusters by three of the four children might help to pinpoint the error locus a little further (see 4.2.2 for the concrete predictions). As discussed above, /sC/ clusters are phonologically different from /C/ + liquid clusters in that /s/ occupies a specific prosodic position, namely the extra-syllabic position (ESP: Kager & Zonneveld, 1986; Fikkert, 1994). The two cluster types show a different timing in development, whereby target /sC/ clusters are produced correctly either later (usually) or earlier than target C + liquid clusters. This seems to imply a difference in processing at the phonological encoding level, rather than a difference in the way segmental information of these cluster types is stored in the mental lexicon: the ESP is either available for the phonological encoding of /s/ or not (yet), just like a complex onset is either available for the phonological encoding of consonants or not. Alternatively, if the absence of a prosodic position in the (developmental) phonological grammar would also affect the linguistic perception of segments in this position, then this could in turn affect the quality of the lexical representation. In this case it remains a problem to pinpoint the exact error locus. Moreover, in this case even a perception experiment would not be able to disambiguate the two possible error loci.

In the final recordings, performance at the PN task – and WR task in case the lexical route is taken there - has improved, and now all three tasks show a similar, but not perfect, performance. It thus appears that the specific problems at the higher levels of lexical access or phonological encoding are solved; it could be the case, for example, that segments can at this point be encoded in syllables with a complex onset. Den Ouden (2002) hypothesizes that if there is a problem at the level of phonetic encoding, all production tasks are affected. Therefore, since performance is still not perfect, and all tasks are now similarly affected, it can be concluded that the main error locus has shifted to the level of phonetic encoding.

Lars exhibits poor performance on all tasks, and this would entail that in his case the problem lies at the level of phonetic encoding, too. This conclusion seems odd in the light of the above discussion of the other children's development. Lars' productions are clearly far more immature than those of the other children, yet the conclusion is that the error locus in his production mechanism is the same as the one for the other children, while this was actually seen as a development from an earlier stage. Two other observations lead me to conclude that in the case of Lars, the entire production mechanism is not functioning well – at least not with respect to the processing of consonant clusters. For one thing, his performance is similar on all tasks, but contrary to the performance of the other children in the final recordings, it is also extremely poor. Secondly, an advantage for the NWR task for target words with fricative-initial clusters (but not sP) is found, where in some cases he produces a single onset consonant instead of omitting the entire cluster. This shows that in addition to problems at the level of phonetic encoding, there are problems at the higher levels of the model, too.

Overall, earlier cluster realizations were found in the NWR task, and the first correct cluster realizations were usually found in this task too. At first sight this appears to be a surprising finding, given that in the study by Hoff et al. (2008), children of the same age usually scored worse on their NWR than on their WR tasks. Hoff et al. state that this suggests that phonological memory demands for the repetition of real words are lower than for repeating non-words. This, in turn, is because the presence of a representational system, like a segmental representation in the mental lexicon, boosts memory capacity by proving an encoding system for the things that need to be remembered. The different findings in the two studies might be related to the different topics of interest, namely in the overall accuracy, measured by PCC in the Hoff et al. study, and in the cluster production ability in the present study. The overall accuracy was not taken into account in the present study, and this could have been lower for the non-words. However, in the study by Vance et al. (2005), where children between the ages of 3-7 years were tested on their performance on PN, WR and NWR tasks, the three-year-olds showed a higher percentage of correctly produced items in both the NWR and the WR repetition task than in the PN task, while the older age groups performed better in both the PN and the WR tasks than in the NWR task. For the three-year-olds the overall results on the WR task were slightly better than those for the NWR task – 66.67% (WR) vs 64.38% (NWR) – but this difference was not significant ($p = 0.289$). For one thing, this suggests that the three-year-olds in the Vance et al. study took the non-lexical route in the WR task, unifying WR and NWR results, while the older children took the lexical route, unifying WR and PN results. Two of the four children in the present study appeared to take the non-lexical route in the WR task as well, while for one the productions in the WR and the PN tasks were very similar, pointing to a lexical route for WR.

Interestingly, for the most complex condition in Vance et al., namely three-syllable words, the NWR had the highest score for the three-year-olds, 58.2%

NWR vs 56.7% WR. This is not a significant difference, but it shows a parallel with the present study, namely that the NWR condition appears to have an advantage in phonologically complex situations. A further parallel between the two studies lies in the fact that the non-lexical route leads to more advanced productions than the lexical route for the three-year-olds. Neither of these parallels can be drawn between the present study and the Hoff et al. study. With respect to phonological complexity, in their first experiment the stimuli in the WR and NWR tasks had the same number of syllables, but of the 9 NWR items 4 started with an onset cluster, compared to only 1 in the WR task – and these were different clusters. There is, thus, not a comparable complex condition in the WR and NWR tasks. In their second experiment they improved the comparability between the two stimuli sets, by replacing all the items with clusters by items with singleton consonants. Of the twelve items per condition, 3 contained three syllables, and performance on these stimuli could present the comparison. However, unfortunately performance on the individual items was not presented in the article, and because the three-syllable words only formed a small part of the total set of test items, differences that might have been there did obviously not survive in the overall performance. With respect to the lexical versus non-lexical route, there was no PN task in the Hoff et al. study, so it cannot be determined whether the participants in their study took the lexical or the non-lexical route in the WR task. In addition to the obvious difference between the small-scale longitudinal study presented here and the single-session, N=15 and N=21 studies of Hoff et al., the differences just discussed make it hard to compare the results of the two studies. The diverging results might thus be due to different underlying factors.

4.6. Conclusion

On the basis of the longitudinal study on the production of onset clusters, I concluded that for three of the four children, Meike, Matteo and Hannah, the initial difficulty in their speech production mechanism lay either in the lexical representation, which could be incomplete, or in the mapping between the phonological and the phonetic encoding level. This conclusion was drawn on the basis of the discrepancy in performance on the NWR and the PN tasks, where in the NWR task more advanced productions were encountered than in the PN task. In the course of time performance on the three tasks became more similar, and improved, indicating a shift in the main error locus from the lexical or phonological encoding level to the phonetic encoding level. In the case of Lars, severe problems with all tasks in addition to evidence for additional problems at the lexical or phonological encoding level led me to conclude that in his case it was impossible to determine exactly what the issues were..

The present study leads to new research questions that should be tested in a carefully set-up, larger study.

A first hypothesis concerns the relative success on production tasks requiring either the lexical route or the non-lexical route. The first adult-like cluster realizations are visible in the NWR task. This points to an initial production advantage for forms that lack a lexical representation over forms that probably have an incomplete or faulty representation in the mental lexicon. This advantage disappears as the segmental representations of real words become more complete and stable in the course of development. From then on, production benefits from this representation; performance on the PN task is no longer worse than performance on the NWR task, and for some children performance on this task even becomes better than the performance on the NWR task. The quality of segmental representations thus seems to be an important factor in the relative performance on production tasks that require taking the lexical route through the model and those that do not require this. As long as representations are unstable or incomplete, the young speaker can excel at repetition tasks, while a stable and complete representation can boost performance on real word tasks, by providing the segments to be produced, together with the established links between the segmental representation and a motor pattern. In the present study, the children's last sessions showed comparable scores on all three tasks. At this point the children had named or repeated the words and the nonwords several times already, and they could have developed motor programs for these words. The presence of a motor program entails that less contextual adaptation of the different phones is required, which facilitates speech production (Nijland & Maassen, 2005). This could have obscured differences between the three tasks in the final recording. This hypothesis about the relative performance on PN and (N)WR tasks and the role of the lexical representation should be studied by carefully balancing real words with and without well-established lexical representations, and systematically comparing the naming of these words to the repetition of similar non-words.

Another question that unfortunately could not be answered in the present study is whether the problem at the initial stages lies with lexical access/the lexical representation or with phonological encoding. This can only be disambiguated by probing the child's lexical representation, for which a well-designed perception experiment is needed (see 4.2.2 for the concrete predictions for a perception-production exmeriment). This is challenging, especially in a longitudinal study. Up until now perception studies with young children have always been group studies, while in this case one would need to compare an individual's performance on perception and production tasks. In addition, in a longitudinal study the challenge is to perform a series of different perception tasks, in order to avoid too much task-experience or boredom.

Appendix 1: Transcriptions of the words and nonwords in Meike's onset cluster development in three production tasks over time (1-19); transcriptions of real words used in the PN and the WR tasks (20 - 23).

1. /Cl/ words: the development of the nonword *bliep* and the word *bloem*.

	words	Sess1	Sess2	Sess4	Sess5	Sess6
NWR	*bliep*	blip	blip	blip	blip	blɪp
PN	*bloem*	blumə	blum	plum	blumə	blum
WR	*bloem*	plum	blum	blumə	blumə	blum

2. /Cl/ words: the development of the nonword *klot* and the word *klok*.

	words	Sess1	Sess2	Sess4	Sess5	Sess6
NWR	*klot*	kle:s	klɔk	klɔt	lɔt	klɔt
PN	*klok*	klɔk	klɔk	klɔk	klɔk	
WR	*klok*	lɔk	klɔk	klɔk	klɔk	klɔk

3. /Cr/ words: the development of the nonword *traak* and the word *trein*.

	words	Sess1	Sess2	Sess4	Sess5	Sess6
NWR	*traak*	kra:s	kra:k	tra:k	tra:k	tra:k
PN	*trein*	trɛin	trɛin	trɛin	trɛm	trɛin
WR	*trein*	trɛin	trɛin	tərɛin	rɛin	trɛin

4. /Cr/ words: the development of the nonword *droon* and the word *draakje*.

	words	Sess1	Sess2	Sess4	Sess5	Sess6
NWR	*droon*		dron	dro:n	dro:m	dro:n
PN	*draakje*	ra:kjə	xra:tjə	dra:k	dra:kjə	

WR	*draakje*	tra:kjə	dra:tjə	dra:kjə	dra:kjə	dra:kjə

5. /Cr/ words: the development of the nonword *kriep* and the word *kraan*.

	words	Sess1	Sess2	Sess4	Sess5	Sess6
NWR	*kriep*	krɛs	krip	krip	kip	klip
PN	*kraan*	kra:n	krɛin	tra:n	kra:n	
WR	*kraan*	kra:n	krɛin		tra:m	kra:n

6. /Cr/ words: the development of the nonword *braak* and the word *broek*.

	words	Sess1	Sess2	Sess4	Sess5	Sess6
NWR	*braak*	bra:s	bra:k	bra:k	bra:k	
PN	*broek*	brukjə	bruk	brukjə	bruk	
WR	*broek*		bruk		bruk	bruk

7. /fric+r/ words: the development of the nonword *friep* and the word *fruit.*

	words	Sess1	Sess2	Sess4	Sess5	Sess6
NWR	*friep*	frik	frip	fwip	frip	frip
PN	*fruit*	frœʏt	frœʏt	frœʏt	frœʏt	frœʏt
WR	*fruit*	prœʏt	frœʏt	frœʏt	frœʏt	frœʏt

8. /fric+r/ words: the development of the nonword *graak* and the word *gras.*

	words	Sess1	Sess2	Sess4	Sess5	Sess6
NWR	*graak*	xra:k	xra:	xra:k	xra:k	xra:k
PN	*gras*	xɑs	xrɑʦ	xrɑs	xrɑs	xrɑs
WR	*gras*	xrɑs	xɑs		xɑs	xrɑs

9. /kn/ words: the development of the nonword *knaak* and the word *knoopjes*

	words	Sess1	Sess2	Sess4	Sess5	Sess6
NWR	*knaak*	knɪk	kla:k	kna:k	xna:k	kna:k
PN	*knoop*	klo:pəs	klo:pjəs	klo:pjəs	kno:pjəs	kno:pjəs
WR	*knoopjes*	klo:pjes	klo:pjəs	no:pjəs	kno:pjəs	kno:pjəs

10. /s+fric/ word: the development of the nonword *schaag* and the word *schaap.*

	words	Sess1	Sess2	Sess4	Sess5	Sess6
NWR	*schaag*	xɑk	xa:x	xa:x	xa:x	xa:x
PN	*schaap*	xa:pjə	xa:p	xa:p	xra:p	
WR	*schaap*	xa:p	xa:p	xa:p	xa:p	

11. /sC/ words: the development of nonword *skaam* and the word *skippybal.*

	words	Sess1	Sess2	Sess4	Sess5	Sess6
NWR	*skaam*	ka:n	ka:m	ka:n	ka:n	ka:m
PN	*skippybal*	kɪpəbɑl	kɪpibɑl	kɪpibɑl	kipɪbɑl	kipɪbɑl
WR	*skippybal*		kɪpibɑl	pɪpibɑ	kipɪbɑl	kipɪbɑl

12. /sC/ words: the development of nonword *spaam* and the word *speeltuin.*

	words	Sess1	Sess2	Sess4	Sess5	Sess6
NWR	*spaam*	pa:s	pa:n		pa:n	spa:m
PN	*speeltuin*	pe:lətœʏn	pe:lpɛin	pe:ltœʏn	pe:ltœʏn	pe:ltœʏn
WR	*speeltuin*	pe:tœʏn	pe:ltœʏn	be:ltœʏn	be:ltœʏn	pe:ltœʏn

13. /sn/ word: the development of nonword *snaak* and the word *snoep.*

	words	Sess1	Sess2	Sess4	Sess5	Sess6
NWR	*snaak*	se:k	sna:k	na:k	sna:k	sna:k
PN	*snoep*	sup	ʃupjsə	knupjəs	fnupjəs	snup
WR	*snoep*	tnup	sup			snup

14. /tv/ word: the development of nonword *twot* and the word *twee.*

	words	Sess1	Sess2	Sess4	Sess5	Sess6
NWR	*twot*				vɔt	tvɔt
PN	*twee*		ve:	dve:	ve:	fve:
WR	*twee*	tve:	fe:	ve:	ve:	fve:

15. /fric+l/ words: the development of nonword *gler* and the word *glas*.

	words	Sess1	Sess2	Sess4	Sess5	Sess6
NWR	*gler*	xle:t	xɛpla	xəlɛr	xjɛl	klɛr
PN	*glas*	xlɑs	xlɑs	las	xlɑs	xlɑs
WR	*glas*	lɑs	xlɑs	xlɑs	klɑs	xlɑs

16. /fric+l/ words: the development of the nonword *vloon* and the word *vlinder*.

	words	Sess1	Sess2	Sess4	Sess5	Sess6
NWR	*vloon*	fle:t	flo:n	flo:n	flo:n	flo:n
PN	*vlinder*		flɪnəɹ	flɪnər		
WR	*vlinder*		flɪŋər	flɪndər	flɪnə	flɪndər

17. /fric+l/ words: the development of the nonword *flaak* and the word *flesje*.

	words	Sess1	Sess2	Sess4	Sess5	Sess6
NWR	*flaak*	flɑp	fla:k	la:k	fla:k	kla:k

PN	*fles*	flɛsjə	flɛsjə	flɛsjə	flɛsjə	
WR	*flesje*	lɛs	flɛsjə	flɛsjə	flɛsjə	flɛs

18. /fric+l/ words: the development of nonword *sloon* and the word *slingers.*

	words	Sess1	Sess2	Sess4	Sess5	Sess6
NWR	*sloon*	plo:n			slo:n	flo:n
PN	*slingers*	siliŋə	tlɪŋərs	slɪŋəs	tlɪŋərs	slɪŋərs
WR	*slingers*		tlɪŋərs		slɪŋərs	slɪŋərs

19. /s+fric/ words: the development of the nonword *zwiep* and the word *zwart*.

	words	Sess1	Sess2	Sess4	Sess5	Sess6
NWR	*zwiep*					
PN	*zwart*	twɑrt	vɑrt		vɑt	
WR	*zwart*	zʊat	fɑrt		vɑt	vɑrt

20. /s+fric/ words: the development of the word *zwembad* in the WR and the PN tasks.

	words	Sess1	Sess2	Sess4	Sess5	Sess6
PN	*zwembad*	sfɛmbɑt	fɛmbɑt	vɛmbɑt	vɛmbɑt	
WR	*zwembad*	pəpɑt	pɛmbɑt	vɛmbɑt	fɛmbɑt	

21. /Cr/ words: the development of the word *kroon* in the WR and the PN tasks.

	words	Sess1	Sess2	Sess4	Sess5	Sess6
PN	*kroon*		xro:n	kro:n	kro:ntjə	
WR	*kroon*		kro:n	xo:n	kro:n	kro:n

22. /s+fric/ words: the development of the word *schaar* in the WR and the PN tasks.

	words	Sess1	Sess2	Sess4	Sess5	Sess6
PN	*schaar*		xa:r		xra:rs	
WR	*schaar*		xa:r	xa:r	xa:r	

23. /s+fric/ words: the development of the word *schoen* in the WR and the PN tasks.

	words	Sess1	Sess2	Sess4	Sess5	Sess6
PN	*schoen*	xun	xun	xun	xrɔn	
WR	*schoen*		xun	xun	xuns	xun

Appendix 2: Transcriptions of the words and nonwords in Matteo's onset cluster development in three production tasks over time (1-19); transcriptions of real words used in the PN and the WR tasks (20 - 23).

1. /Cl/ words: the development of the nonword *bliep* and the word *bloem*.

	words	Sess1	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*bliep*	pip	blip	blip	blip	blip	blip
PN	*bloem*	obumə	pum	bum	blumə	bluməʧə	blumə
WR	*bloem*	bum	bum	bumə	bumə	blumə	blumə

2. /Cl/ words: the development of the nonword *klot* and the word *klok*.

	words	Sess1	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*klot*	kɔp	kɔps	kɔk	klɔt	klɔt	klɔk
PN	*klok*		kɔk	kɔk	klɔk	klɔk	klɔk
WR	*klok*	kɔk	kɔk	kɔk	klɔk	klɔt	klɔk

3. /Cr/ words: the development of the nonword *braak* and the word *broek*.

	words	Sess1	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*braak*		ba:k	ba:k	bla:k	bra:k	bra:k
PN	*broek*	puk	buk	buk	bluk	bruk	bruk
WR	*broek*	buk	puk		bluk	bluk	bruk

4. /Cr/ words: the development of the nonword *kriep* and the word *kraan*.

	words	Sess1	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*kriep*	kɪp		kɪp	kəip	kɪp	krip
PN	*kraan*	ta:n	ka:n	ka:n	kla:n	kra:n	kra:n
WR	*kraan*	ta:n	ka:n	tla:n	kla:n	kro:n	kra:n

5. /Cr/ words: the development of the nonword *traak* and the word *trein*.

	words	Sess1	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*traak*	ka:k			tra:k	dra:k	tra:k
PN	*trein*		tɛin	tɛin	ʦɛin	trɛin	trɛin
WR	*trein*	trɛin	pɛi	tɛin	tlɛin	trɛin	trɛin

6. /Cr/ words: the development of the nonword *droon* and the word *draak*.

	words	Sess1	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*droon*	bo:m		do:m	dlo:n	dlo:n	dro:n
PN	*draakje*	ta:tjə	ta:tjə	da:kjə	tla:k	da:kjə	dra:kjə
WR	*draakje*	ta:kjə	da:kjə	ta:kjə	tra:kjə	dla:kjə	dra:kjə

7. /fric+r/ words: the development of the nonword *friep* and the word *fruit*.

	words	Sess1	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*friep*	tip	fip	tlip	flip	flɪp	flip
PN	*fruit*	rœit	fœʏt	fœʏt	frœʏt		frawtə
WR	*fruit*	tœit	fœʏt	fœʏs	flœʏt	fwœʏt	frœʏt

8. /fric+r/ words: the development of the nonword *graak* and the word *gras*.

	words	Sess1	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*graak*	ka:k	ka:k	xra:k	xa:k	xla:k	xra:k
PN	*gras*	xɑʦ	xɑs	xɑs	xlɑs	xlɑs	xrɑs
WR	*gras*	kɑs		xrɑs		klɑs	xrɑs

9. /kn/ words: the development of the nonword *knaak* and the word *knoop*.

	words	Sess1	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*knaak*	kakok	ka:k	ka:	kla:k	kra:k	kna:k
PN	*knoop*	ko:pjəs	ko:p	ko:pjəs	klo:pjəs	klo:pjəs	kno:pjəs
WR	*knoopjes*	ko:pi	ko:pjə	o:kjəs	klo:pjəs	klo:pis	kno:pjəs

10. /s+fric/ words: the development of the nonword *schaag* and the word *schaap*.

	words	Sess1	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*schaag*	trax		xa:x	sxa:k	xla:x	sxa:r
PN	*schaap*	ta:p		a:pjə	xa:p	xa:p	sxra:p
WR	*schaap*	ka:p	xa:p	xa:	xa:p	sxa:pjə	sxa:p

11. /sC/ words: the development of nonword *spaam* and the word *speeltuin*.

	words	Sess1	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*spaam*	əspa:m		spa:m	spam	spa:m	sfpa:m
PN	*speeltuin*	pɪtœin	spiltœʏn		spe:ltœʏn	spe:ltœʏn	pe:ltœʏn
WR	*speeltuin*	petœin	spe:ltœʏn	spe:ltœʏn	spe:ltœʏn	spe:ltœʏn	pe:ltœʏn

12. /sC/ words: the development of nonword *skaam* and the word *skippybal*.

	words	Sess1	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*skaam*	kra:n		ka:	ska:m	ska:m	ska:m
PN	*skippybal*	kɪkɪbɑl		kɪpibɑl	skipɪbɑl	skɪpibɑl	skɪpɪbɑl
WR	*skippybal*			hkɪpibɑl	kikɪbɑl	skɪpibɑl	skɪpibɑl

13. /sn/ words: the development of the nonword *snaak* and the word *snoep*.

	words	Sess1	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*snaak*	ka:k		tfa:k	sla:k	sla:k	sna:k
PN	*snoep*	tupis		supjəs	supjəs	supjəs	snupjəs
WR	*snoep*	sup	sup	su:p	slup	sup	snup

14. /tv/ words: the development of the nonword *twot* and the word *twee*.

	words	Sess1	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*twot*		sɔpf	sɔt	tvɔt	tɔt	prɔt
PN	*twee*	te:	pe:	te:	tve:	tre:	tre:
WR	*twee*	te:	pe:	pɛ:	tve:	tve:	tve:

15. /fric+l/ words: the development of the nonword *gler* and the word *glas*.

	words	Sess1	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*gler*	sɛr	xrə	xlə	klɛm	xlɛr	xlɛl
PN	*glas*	aʃjə	χɑ:s	xa:	klɑs	xlɑs	xlɑs
WR	*glas*	kɑs		khɑs	klɑ	xlɑs	klɑs

16. /fric+l/ words: the development of the nonword *floon* and the word *vlinder*.

	words	Sess1	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*floon*	fo:m	fo:n	fwo:m	flo:n	flo:n	flo:n
PN	*vlinder*	tɪn:	sɪn:ə	fɪndəɹ	fɪn:əɹ	flɪnə	
WR	*vlinder*	fɪnə	sɪndəɹ	fɪnə	flɪnən	xlɪndəɹ	flɪndər

17. /fric+l/ words: the development of the nonword *flaak* and the word *fles*.

	words	Sess1	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*flaak*	ka:t		fa:k	fla:k	fla:k	fla:k
PN	*fles*		fɛs	fɛs	flɛs	flɛs	flɛs
WR	*fles*		fɛsjə	fɛsjə	flɛʃjə	flɛsjə	flɛsjə

18. /fric+l/ words: the development of the nonword *sloon* and the word *slingers*.

	words	Sess1	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*sloon*	xowom		so:n	slo:m	slo:k	slo:m
PN	*slingers*	tɪnəs	slɪn:ə	sɪn:əs	slɪndəs	slɪnəs	slɪnəs
WR	*slingers*	tɪtəs		sɪn:əs	slɪn:əs	slɪ:əs	slɪndəs

19. /s+fric/ words: the development of the nonword *zwiep* and the word *zwart*.

	words	Sess1	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*zwiep*	plip					zʊip
PN	*zwart*	tɑs	sɑt	səɑ:t	zwɑt	zwɑt	zʊɑrt
WR	*zwart*	sɑt	sɑt	nɑt	zʊɑʦ	vɑt	zʊɑt

20. /fric+v/ words: the development of the word *zwembad* in the WR and the PN tasks.

	words	Sess1	Sess3	Sess4	Sess5	Sess6	Sess7
PN	*zwembad*	pɛmtɑt	sɛmbɑt	sɛm:ɑt	stsɛmbɑt	sɛmbɑt	zʊɛmbɑt
WR	*zwembad*	sɛmbɑt	sɛm:ɑt	sɛm:ɑt	sɛmbɑt	slɛmbɑt	sʊɛmbɑt

21. /Cr/ words: the development of the word *kroon* in the WR and the PN tasks.

	words	Sess1	Sess3	Sess4	Sess5	Sess6	Sess7
PN	*kroon*	ko:m	ko:ntjə		klo:m	kro:n	kro:n
WR	*kroon*	xo:n	ko:m	klo:n	kro:m	kro:n	kro:n

22. /s+fric/ words: the development of the word *schoen* in the WR and the PN tasks.

	words	Sess1	Sess3	Sess4	Sess5	Sess6	Sess7
PN	*schoen*	ku:n	χunə	xun	xun	xun	sxun
WR	*schoen*	xun	xun	kun	kxun	sun	sxun

23. /s+fric/ words: the development of the word *schaar* in the WR and the PN tasks.

	words	Sess1	Sess3	Sess4	Sess5	Sess6	Sess7
PN	*schaar*	xa:r	xa:ɹ	xa:	sxa:t	sxa:r	sxa:t
WR	*schaar*		χa:ɹ	xa:r	fa:r	xa:r	sxa:r

Appendix 3: Transcriptions of the words and nonwords in Hannah's onset cluster development in three production tasks over time (1-19); transcriptions of real words used in the PN and the WR tasks (20 - 23).

1. /Cl/ words: the development of the nonword *bliep* and the word *bloem*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*bliep*	pip	klip	blip	kip		blip	blip
PN	*bloem*	bumə	pumətjə	bumətjə	blumə	pɔmətjəs	blumətjəs	blum
WR	*bloem*	bumə		bəlumə		blumə	blumə	blum

2. /Cl/ words: the development of the nonword *klot* and the word *klok*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*klot*	klɔs	klɔt	klɔt	klɔt	klɔt	klɔt	klɔt
PN	*klok*		tlɔt	klɔt	klɔt	klɔt	klɔt	klɔt
WR	*klok*	slɔt	klɔt	klɔt	klɔt		klɔt	klɔt

3. /Cr/ words: the development of the nonword *traak* and the word *trein*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*traak*	taːt	daːt	daːt	baːt	taːt	taːt	daːk
PN	*trein*	tɛin	tɛin	tɛitjə	tɛin	stɛin	tɛin	tɛin
WR	*trein*		tɛin	tɛin	tɛin	tɛin	tɛin	tɛin

4. /Cr/ words: the development of the nonword *droon* and the word *draak*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*droon*	toːn	doːn	doːn	doːn	doːn	doːn	doːn
PN	*draakje*	taːtjə		daːtjə	daːjə	daːx		daːk

WR	*draakje*	taːtjə	daːtjə	draːtjə		draːtə	daːtjə	daːtjə

5. /Cr/ words: the development of the nonword *kriep* and the word *kroon*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*kriep*	klaːt		klɪpɪp	pip	tip	dip	dip
PN	*kroon*	toːn	toːn	toːn		toːn		
WR	*kroon*	toːn	toːn					toːn

6. /Cr/ words: the development of the nonword *braak* and the word *broek*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*braak*	klaːt		baːt		daːt	daːt	daːt
PN	*broek*	putjə	put	put	buts	but	but	but
WR	*broek*	put	but		but	buts	but	bup

7. /fric+r/ words: the development of the nonword *friep* and the word *fruit*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*friep*	pip	pjip	pip		flip	fip	pɪp
PN	*fruit*	poeʏt	poeʏt	flœʏt	fœʏt		foeʏt	poeʏt
WR	*fruit*	poeʏt			foeʏt	foeʏt	foeʏt	poeʏt

8. /fric+r/ words: the development of the nonword *graak* and the word *gras*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*graak*	tlaːt		xarə	xraːk	xaːt	xəlaːt	xaːt
PN	*gras*	tɑs	xlɑs	xɑs		xɑs	xɑs	xlɑs
WR	*gras*		klɑs	xrɑs	glɑs	xɑs	xɑrs	xɑs

9. /kn/ words: the development of the nonword *knaak* and the word *knoop*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*knaak*	knaːt	taːk	taːt		tənaːt	kənaː	knaːk
PN	*knoopjes*	poːmpjəs	poːntjəs	boːmpjəs	poːpjəs	tənoːpjəs	təloːpjəs	tənoːpjəs
WR	*knoopjes*	pompəs	poːmpjs		poːpjə	tənoːpjəs	tənoːpjəs	tənoːpjəs

10. /s+fric/ words: the development of the nonword *schaag* and the word *schoen*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*schaag*	tlaːx	taːx		saːx	staːn	staː	staːt
PN	*schaap*	paːpjə		paːpjə	paːpjə	staːp	staːp	staːp
WR	*schaap*		paːp	paːpjə	taːp	staːp	staːp	

11. /sC/ words: the development of nonword *skaam* and the word *skippybal*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*skaam*	taːm	kaːm		taːm		staːm	staːn
PN	*skippybal*	pitɪbɑl	pipɪbɑl	pipɪbɑl	papɛbɑl	pipɪbɑl	stitɪbɑl	tipɪbɑl
WR	*skippybal*	pipɪbɑl	pipɪbɑl	pipɪbɑl	pipɪbɑl	spipɪbɑl	stipɪbɑl	stibɑl

12. /sC/ words: the development of nonword *spaam* and the word *speeltuin*.

	words	Sess1	Sess2	Sess3	Sess4	Sess6	Sess7
NWR	*spaam*	spaːn	paːm	spaːm	paːn	spaːm	paːm
PN	*speeltuin*		putoeʏn	peːltoeʏn	pitœʏn		speːltoeʏn
WR	*speeltuin*	peːn	peːltoeʏn		peːltœʏn	speːltoeʏn	speːltoeʏn

13. /sn/ words: the development of the nonword *snaak* and the word *snoep*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*snaak*	kla:t			bra:t	sna:t	tna:k	xna:t
PN	*snoep*	pupjəs	pumpjəs	pupjəs	do:pjəs	snupjəs	snupjəs	snup
WR	*snoep*	pup		pumpjəs	stənup	snup	snup	snup

14. /tv/ words: the development of the nonword *twot* and the word *twee*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*twot*	pɔt	pɔt	vɔt	tlɔt	tɔt	tɔt	dvɔt
PN	*twee*	teː	teː	teː			tʊeː	teː
WR	*twee*	teː	teː	teː	teː	teː		tveː

15. /fric+l/ words: the development of the nonword *gler* and the word *glas*.

	words	Sess1	Sess2	Sess4	Sess5	Sess6	Sess7
NWR	*gler*	glɛt	klɑp		plɛ	xlɛp	xəlɛr
PN	*glas*	xlɑs		xlɑs	xlɑs	xlɑs	xɑs
WR	*glas*	tlɑs	xlɑs	slɛs	xlɑs	xlɑs	xlɑs

16. /fric+l/ words: the development of the nonword *floon* and the word *vlinder*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*floon*	poːn		foːn	floːn	floːn	floːn	foːn
PN	*vlinder*	fɪnəɹ	pɪndər		slɪndər	flɪnətjə	flɪnər	pɪndə
WR	*vlinder*	pɪnə	fɪndər	pɪndər	flɪnɛr	flɪnə	flɪnər	pɪnə

17. /fric+l/ words: the development of the nonword *flaak* and the word *flesje*.

NWR	*flaak*	falabak	flaː		flaːk	flaːk	flaːt	flaːt
WR	*fles*	fεs	pεsjə	fεsjə		flεs	flεs	flεs
PN	*flesje*	pεtjə	pεsjə	flεsjə		flεsjə	flεs	flεs

18. /fric+l/ words: the development of the nonword *sloon* and the word *slingers*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*sloon*	kloːn	sloːn	tloːn	tloːn	sloːn		sloːn
PN	*slingers*	tlɪnəs	tlɪnə			slɪnəs	səlɪnə	slɪnəs
WR	*slingers*	tlɪndəs	klɪkjəs	klɪnəs	tlɪnəs	slɪnərs	sxlɪnəs	slɪnə

19. /s+fric/ words: the development of the nonword *zwiep* and the word *zwart*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*zwiep*	pip						
WR	*zwart*	fɑrt	pɑrt	pɑrt	bɑːt	pɑrt	zpɑt	
PN	*zwart*	pɑrt	fɑrt	pɑrt	pɑrt	pɑrt	zʊɑt	zʊɑt

20. /s+fric/ words: the development of the word *zwembad* in the WR and the PN tasks.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
WR	*zwembad*	pεmpɑt	pεmbɑt	pεmbɑt	sεmbɑt	sεmbɑt	sεmbɑt	stεmbɑt
PN	*zwembad*	pεmbɑt		pεmbɑt	sεmbɑt		sεmbɑt	sεmbɑt

21. /Cr/ words: the development of the word *kroon* in the WR and the PN tasks.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
WR	*kroon*	toːn	toːn					toːn
PN	*kroon*	toːn	toːn	toːn		toːn		

22. /s+fric/ words: the development of the word *schaar* in the WR and the PN tasks.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
WR	*schaar*	taːr	taːr	taːr		staːr	staːr	staːr
PN	*schaar*	taːr	tsaːr				staːr	staːr

23. /s+fric/ words: the development of the word *schoen* in the WR and the PN tasks.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
WR	*schoen*	tun	tun	tun	tunə	stun	stun	stun
PN	*schoen*	tunə	tunə	tunə	tunə		stun	stun

Appendix 4: Transcriptions of the words and nonwords in Lars' onset cluster development in three production tasks over time (1-19); transcriptions of real words used in the PN and the WR tasks (20 - 23).

1. /Cl/ words: the development of the nonword *bliep* and the word *bloem*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*bliep*	bits	bip			bif	bip	bipf
PN	*bloem*	bo			bum:	bun	bum	bumə
WR	*bloem*	bəl		bubu	bumə	bun		bumə

	words	Sess8	Sess9	Sess10	Sess11
NWR	*bliep*	bip	tɪp	blip	pis
PN	*bloem*	bumə	bum	bum	bum ma
WR	*bloem*	bumə	a bumə	bumə	pumə

2. /Cl/ words: the development of the nonword *klot* and the word *klok*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*klot*	pɔts		tɔts	dɔts	tɔts	tɔts	
PN	*klok*	dɔt		dɔt	dɔts	tɔts	dɔts	dɔts
WR	*klok*	tɔts			tɔts	dɔts		tɔts

	words	Sess8	Sess9	Sess10	Sess11
NWR	*klot*	tɔt	tɔts	tɔts	tɔts
PN	*klok*	tɔts	tɔts	tɔts	tɔɹts
WR	*klok*	tɔts	a tɔts	tɔts	tɔts

3. /Cr/ words: the development of the nonword *traak* and the word *trein*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*traak*	dado		taːts	baːt	taːts	taːts	taːt
PN	*trein*	tɛi	tɛi	tɛi	tɛi	tɛi	tʀɛin	tɛin
WR	*trein*	tɛi	tɛi	tɛit	tɛjə	tɛin		tɛin

	words	Sess8	Sess9	Sess10	Sess11
NWR	*traak*		taːts	taːs	taːts
PN	*trein*	tɛin		tɛin	tɛin
WR	*trein*	tɛin	tɛin	tɹɛin	tɛin

4. /Cr/ words: the development of the nonword *droon* and the word *draak*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*droon*	toːn	doːn	doːn	doːn	doːn	doːn	doːn
PN	*draakje*	taːtjə		daːtjə	daːjə	daːx		daːk
WR	*draakje*	taːtjə	daːtjə	draːtjə		draːtə	daːtjə	daːtjə

	words	Sess8	Sess9	Sess10	Sess11
NWR	*droon*			doːn	toːn
PN	*draakje*	taːɹt	taːtjə	taːts	
WR	*draakje*	taːtsə	daːtjə	taːtsjə	taːsjɛ

5. /Cr/ words: the development of the nonword *kriep* and the word *kroon*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*kriep*	tɪts			dip		pip	bit
PN	*kroon*	doː		toːn	taŋə	woːn	doːn	toːn
WR	*kroon*	toː			toːn	boːm		toːn

	words	Sess8	Sess9	Sess10	Sess11
NWR	*kriep*	bip	a pip	pip	pɪp
PN	*kroon*	toːn	toːn	toːn	toːn
WR	*kroon*	toːn	toːn	toːn	toːnɹ

6. /Cr/ words: the development of the nonword *braak* and the word *broek*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*braak*	bats	baːts		bae	taːts	taːts	taːts
PN	*broek*	buts	buts	buts	puts	buts	buts	buts
WR	*broek*			buts	buts	buts	buts	buts

	words	Sess8	Sess9	Sess10	Sess11
NWR	*braak*	taːts	taːts	taːts	paːɹts
PN	*broek*		buts	buts	puts
WR	*broek*	buts	buts	buts	

7. /fric+r/ words: the development of the nonword *friep* and the word *fruit*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*friep*	ɪp	ip		its	ɪf	ip	fip
PN	*fruit*	œʏt	œʏt	œʏt	œʏt	œʏt	œʏts	œʏt
WR	*fruit*		atʃ		œʏt	œʏts	œʏts	tœʏt

	words	Sess8	Sess9	Sess10	Sess11
NWR	*friep*	ip	ɪp	ip	ip
PN	*fruit*	œʏt	œʏt	kœʏ	œʏt
WR	*fruit*	œʏt	œʏt	œʏʃ	œʏt

8. /fric+r/ words: the development of the nonword *graak* and the word *gras*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*graak*	ats			aːts	dɑs	aːts	aːts
PN	*gras*	ɑs	ɑs	ɑs	ɑs	ɑs	ɑs	ɑs
WR	*gras*	ɑs		ɑs	ɑs	ɑs		ɑs

	words	Sess8	Sess9	Sess10	Sess11
NWR	*graak*	aːts	aːts	aːs	aːs
PN	*gras*	ɑs	ɑs		ɑs
WR	*gras*	ɑɹs	ɑs	ɑs	ɑɹs

9. /kn/ words: the development of the nonword *knaak* and the word *knoop*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*knaak*	taːts	kæːk		taːts	taːts	taːts	taːts
PN	*knoopjes*	butə		doːtjə	dotsjə	doːtʃə	toːtʃə	boːtʃə
WR	*knoopjes*	butə		doːtjə	dotsjə	doːtʃə	toːtʃə	boːtʃə

	words	Sess8	Sess9	Sess10	Sess11
NWR	*knaak*	taːts	taːts	taːs	taːts
PN	*knoopjes*	toːtʃəs	poːmptəs	poːtjəs	boːtəs
WR	*knoopjes*	toːtʃəs	poːmptəs	poːtjəs	boːtəs

10. /s+fric/ words: the development of the nonword *schaag* and the word *schoen*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*schaag*	aːp	aːp	doː	ae	asː	s aːs	aːts
PN	*schaap*		aːp		aːpf	aːp	s aːp	aːp
WR	*schaap*			aːp	aː	aːp		aːpf

	words	Sess8	Sess9	Sess10	Sess11
NWR	*schaag*	aːts	aːs	haː	aːs
PN	*schaap*	aːptʃə	aːtsə	aːp	
WR	*schaap*	aːp	baːptsə	ampsə	

11. /sC/ words: the development of nonword *skaam* and the word *skippybal*.

	Words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*skaam*	taːm			tan	daːn	taːm	taːn
PN	*skippybal*	pɪti	pɪpi bɑl	pɪpibɑl	pipi bɑlə	pipibɑl	spipibɑl	bipibɑl
WR	*skippybal*	pɪti		pɪpibɑl	bɪpibɑl	pipibɑl	pipibɑl	

	words	Sess8	Sess9	Sess10	Sess11
NWR	*skaam*	taːn	aːn	taːn	paːn
PN	*skippybal*	pipibɑl		pipibɑl	pipibɑl
WR	*skippybal*	pipibɑl	pipibɑl	bipibɑl	pipibɑl

12. /sC/ words: the development of nonword *spaam* and the word *speeltuin*.

	words	Sess1	Sess2	Sess3	Sess4	Sess6	Sess7
NWR	*spaam*	taːm	paː	baː	taːnː	paːn	taː
PN	*speeltuin*	m tœʏn	bu tœʏn	tutœʏn	titœʏ	deːtœʏn	atœʏn
WR	*speeltuin*	p tœʏn		butœʏ	idœʏn	tɛtœyn	

	words	Sess8	Sess9	Sess10	Sess11
NWR	*spaam*	paːn	paːn		paːn
PN	*speeltuin*	teːtœʏn	teːtœʏn	teːltœʏn	teːtœʏn
WR	*speeltuin*	teːtœʏn	a teːtœʏn	teltœʏn	teːtœʏn

13. /sn/ words: the development of the nonword *snaak* and the word *snoep*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*snaak*	taːts			ɑts	aŋ	aːts	aːhts
PN	*snoep*		ɪpi	dutjə	upjə	uftʃɛ	upjə	uptsə
WR	*snoep*	ɔp			ups	upf		ups

	words	Sess8	Sess9	Sess10	Sess11
NWR	*snaak*	aːts	aːts	aːs	aːts
PN	*snoep*	upjəs	upjəs	utjəs	ba utəɹs
WR	*snoep*		a up		

14. /tv/ words: the development of the nonword *twot* and the word *twee*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*twot*	dɑts	ɔs			tɔts	dɔts	
PN	*twee*				teː	tɛ	teː	teː
WR	*twee*				tse	ti		teː

	words	Sess8	Sess9	Sess10	Sess11
NWR	*twot*		tɔts	tɔts	tɔɹs
PN	*twee*		teː	tjeː	teː
WR	*twee*	teː	teː	teː	teː

15. /fric+l/ words: the development of the nonword *gler* and the word *glas*.

	words	Sess1	Sess2	Sess4	Sess5	Sess6	Sess7
NWR	*gler*	dɛr			eː		ehe
PN	*glas*	ɑs	ɑs	ɑs		ɑs	ɑs
WR	*glas*	ɑs			ɑs	ɑs	

	words	Sess8	Sess9	Sess10	Sess11
NWR	*gler*		ɛɹ	ɛɹ	ɛɹ
PN	*glas*	ɑs	ɑs	ɑs	ɑɹs
WR	*glas*	ɑs	a ɑs	ɑs	aɹs

16. /fric+l/ words: the development of the nonword *floon* and the word *vlinder*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*floon*		fwoː		oːn	oːn	oːn	dɔts
PN	*vlinder*			ɪː	ɪŋəh	ɪnːɛ	ɛnɛ	ɪnːɛ
WR	*vlinder*	ɪ	ɪ		ɪŋə	ɪnːə	ɛnə	ɪnːə

	words	Sess8	Sess9	Sess10	Sess11
NWR	*floon*	oːn	oːn	oːn	oːn
PN	*vlinder*	ɪnə	ɪndəɹ		ɪnəɹ
WR	*vlinder*	ɪnːə	ɪnə	ɪnə	inəɹ

17. /fric+l/ words: the development of the nonword *flaak* and the word *flesje*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*flaak*	ats		aːts	ats	ats	aːts	slaːk
PN	*fles*	ɛs	ɛs	flɛs	ɛs	ɛs	ɛs	
WR	*flesje*	ɛs			ɛsː	ɛs		ɛsɛ

	words	Sess8	Sess9	Sess10	Sess11
NWR	*flaak*	aːts	aːts	aːts	aːts
PN	*fles*	ɛs	ɛs	ɛs	ɛs
WR	*flesje*	esjə	ɛsɛ		ɛsɛ

18. /fric+l/ words: the development of the nonword *sloon* and the word *slingers*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*sloon*		loː		aŋ	hoːn	oːn	oːn
PN	*slingers*			ɪː	ɪntej	ɪŋə	ɛnɛ	ɛnə
WR	*slingers*				ɪndis	ɛnɪs		ɪnəɹ

	words	Sess8	Sess9	Sess10	Sess11
NWR	*sloon*	oːn	oːn	oːn	oːn
PN	*slingers*	ɪnəs	ɪnə	ənəs	
WR	*slingers*	ɪnəs	ɪnəs	ɪnəs	ɪnəs

19. /s+fric/ words: the development of the nonword *zwiep* and the word *zwart*.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
NWR	*zwiep*	ɪp			ɪp	ɪp	ɪp	ɪpf
WR	*zwart*	ɑts		dɑːt	ɑts	aɹt	ɑs	ɑts
PN	*zwart*	ɑts			ats	ɑts	ɑts	tɑts

	words	Sess8	Sess9	Sess10	Sess11
NWR	*zwiep*	ɪp	ɪp	ip	ɪp
PN	*zwart*	ɑts	ɑts	ɑɹts	ɑɹt

WR	*zwart*	ɑts	ɑts	ɑɹts	aɹt

20. /s+fric/ words: the development of the word *zwembad* in the WR and the PN tasks.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
WR	*zwembad*	m bɑts	bɑt	bɛbɑt	ɛ bɑts	bɑts	ɛbasː	ɛbɑts
PN	*zwembad*	bɑts			pɛmpɑt	ɛbɑt		bɑt

	words	Sess8	Sess9	Sess10	Sess11
PN	*zwembad*	ɛbɑts		ɛmbɑt	ɛmbɑt
WR	*zwembad*		a ɛmbɑt	ɛmbɑtə	ɛmpɑt

21. /Cr/ words: the development of the word *kroon* in the WR and the PN tasks.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
WR	*kroon*	doː		toːn	taŋə	woːn	doːn	toːn
PN	*kroon*	toː			toːn	boːm		toːn

	words	Sess8	Sess9	Sess10	Sess11
PN	*kroon*	toːn	toːn	toːn	toːn
WR	*kroon*	toːn	toːn	toːn	toːnɹ

22. /s+fric/ words: the development of the word *schaar* in the WR and the PN tasks.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
WR	*schaar*			aːr	ajə	aːɹ	aɹa	aːɹ
PN	*schaar*				aːɹ	aɹ		aːɹ

	words	Sess8	Sess9	Sess10	Sess11
PN	*schaar*	aːɹ		aːɹ	aːɹ
WR	*schaar*	aːɹ		aːɹ	aːɹ

23. /s+fric/ words: the development of the word *schoen* in the WR and the PN tasks.

	words	Sess1	Sess2	Sess3	Sess4	Sess5	Sess6	Sess7
WR	*schoen*	o	o		un	unː	un	un
PN	*schoen*	u		u	un			un

	words	Sess8	Sess9	Sess10	Sess11
PN	*schoen*	un	un	un	
WR	*schoen*	un	a un	un	un

Appendix 5: Words and nonwords used in the three production tasks (PN, WR and NWR) and their respective averaged log transitional probabilities; Dutch orthography is used for the annotation.

words	averaged log transitional probability	nonwords	averaged log transitional probability
draakje	-1.20	droon	-1.17
kroon	-1.15	krip	-1.28
kraan	-1.13	kraak	-1.15
broek	-1.14	braak	-1.11
trein	-1.21	traak	-1.20
gras	-1.14	graak	-1.16
fruit	-1.36	friep	-1.39
speeltuin	-1.34	spaam	-1.12
skippybal	-1.23	skaam	-1.21
schaap	-1.23	schaag	-1.24
schaar	-1.04		
schoen	-1.24		
zwembad	-1.18	zwiep	-1.42
zwart	-1.12		
snoep	-1.49	snaak	-1.28
klok	-1.23	klot	-1.18
bloem	-1.27	bliep	-1.19
vlinder	-0.90	vloon	-1.18
fles	-1.27	flaak	-1.27
glas	-1.26	gler	-1.22
twee	-1.46	twot	-1.21
knoop	-1.37	knaak	-1.21

Chapter 5. Perception of onset clusters by two-year-olds: the case of /Cl/, /Cr/ and /sC/ clusters[1]

5.1. Introduction

In this chapter I examine how two-year-olds *perceive* words with onset clusters that are either reduced or fully pronounced. As in Chapter 1, the focus is on onset clusters consisting either of consonant + liquid, or of /s/ + consonant. When toddlers start producing such words, they typically reduce the word by omitting one of the cluster consonants. However, as careful analyses of the produced tokens show in Chapter 1, there are acoustic traces of the omitted consonant, but these are unperceivable to the adult ear. This suggests that the apparently omitted segments are in any case present in the segmental representation. Toddlers' cluster realizations are thus reduced during the production process, at lower levels of their speech production mechanism. The evidence for this assumption so far only comes from analyses of child productions. To test the hypothesis that toddlers have detailed representations of clusters, we need to examine infants' perception of (reduced) clusters. This is why, in the current chapter I report on a preferential looking experiment designed to investigate whether two-year-olds are sensitive to reduced onset clusters in perception. Production data from most of the children participating in the experiment were obtained too, to observe the link between production and perception in more detail.

I have limited the work in this thesis to a study of the system behind the production of isolated words, since this is what the developing speakers in this thesis, being between one and two-years old, mostly produce. Within the context of word-production, this study will focus on the - developing - production of word-onset consonant clusters. A typical deviation in early child

[1] This Chapter is partly based on the publication:
Gulian, M, Junge, C. & Levelt, C. (2014). Two-year-olds distinguish snakes from nakes but not trains from tains. *BUCLD 38 Proceedings,* Cascadilla Press.

language productions is the reduction of these clusters to singleton consonants, like in (Dutch) [tɛin] for target *trein* 'train', and [tup] for target *stoep* 'side-walk'. As mentioned above, up until now we only find grammatical accounts of this deviation, in the form of a fixed syllable template, a parameter setting, or a constraint on syllable structure (Fikkert, 1994; Pater & Barlow, 2003; Velleman & Vihman, 2002). A brief discussion of these accounts will follow below in 1.4. However, instead of resulting from a specific grammatical setting, these cluster reductions could also be the outcome of the speech production process, and in the speech production mechanism there are several possible sources for error that could be considered. This is what will be done in this thesis, by studying children's cluster productions in different ways - acoustically, phonologically, and in relation to children's perception of consonant clusters - and analyzing both longitudinal, spontaneous production data, and elicited productions.

To examine the amount of detail that toddlers stored for onset clusters in their mental representations, a preferential looking experiment ('PLP': Golinkoff et. al, 1987; for a recent review see Golinkoff et. al, 2013) was carried out. Swingley and Aslin (2000) modified this paradigm to examine the amount of detail with which words are stored in the infant brain: infants listen to correct pronunciations or to 'mispronunciations' of words corresponding to one of two pictures that they are presented with on a screen. To obtain a mispronunciation, usually one of the target phonemes is replaced by another phoneme, like in the mispronunciation 'vaby' instead of the correct 'baby'. Although even in the mispronunciation condition, infants fixate the correct object above chance, their looking behavior is affected by the way words are produced: Infants typically look longer at a target picture that is named correctly than when it is named with a mispronunciation.

Table 1 presents an overview of studies comparing infants' performance for 'correct pronunciations' versus 'mispronunciations'. Most studies provide evidence showing that infants have detailed word representations: infants are sensitive to mispronunciations for consonants as well as for vowels, in different positions of the word (e.g., in onset, in medial, and in coda position). Infants notice mispronunciations both for well-known words and for recently-learned words (Bailey & Plunkett, 2002; White & Morgan, 2008), which suggests that word-templates are detailed from the beginning of storage (see also Altvater-Mackensen & Mani, 2013). However, infants are not sensitive to all mispronunciations: Detection of mispronunciations is dependent on the identity of the target phoneme (Van der Feest, 2007), and on the overlap in phonological features between target and the substituted phoneme (White & Morgan, 2008). However, the mispronunciation paradigm has not yet been extended to words starting with complex clusters, which is the focus of this thesis. Therefore, the current chapter examines how detailed clusters are stored in the mental lexicon of toddlers' speech production mechanism.

Table 1: Overview of the results of the studies summarized in this section.

Study	**Age and language**	**Method**	**Type of mispronunciation+ Example**	**Results**
Swingley and Aslin (2000)	From 18 to 23 months Dutch	PLP	Onset dog → tog	Words stored in detail
Swingley (2003)	19-month-olds Dutch	PLP	Onset + medial position bal→ dal baby→ bady	Detailed storage of plosives in word onset and word medial position
Swingley (2005)	11-month-olds Dutch	HTPP	Onset + coda teen → peen teen → teem	Detailed storage of plosives in word onset but not in word coda
Mani and Plunkett (2007)	15, 18 and 24-month-olds English	PLP	Onset + vowel bed → bud bed → ped	Detailed storage of both vowels and

				consonants
Bailey and Plunkett (2002)	18, 24-month-olds English	PLP	Onset cup → gup (change in voice) cup → dup (change in place and voice)	Detailed storage of words, no correlation with vocabulary
White and Morgan (2008)	19-month-olds English	PLP	Onset cup → tup (change in place) cup → bup (change in place and voice) cup → vup (change in place, voice, manner)	Detailed storage of consonants in the word onset
Van der Feest (2007)	24-month-olds Dutch	PLP	Onset (change in place and/or voice) poes → boes boom → poom poes → does doos → poos	Detailed storage of consonants in the word onset of the voice feature, asymmetric findings for place.
Mani et al. (2012)	14-month-olds English	ERP	Vowels bed → bid	Detailed storage of vowels
Mills et al. (2004)	14 and 20-month-olds	ERP	Onset dog → bog cat → gat	Consonants are not stored in detail

***HTP (headturn procedure); (S)PLP (sequential) preferential looking paradigm); ERP (event-related potentials)
****Kuh 'cow'; Schaf 'sheep'; Taf (MP of Schaf); Buch 'book'.

When toddlers try to produce onset clusters, they often reduce the cluster by omitting one of the consonants. Which consonant is omitted depends on the identity of the cluster. For consonants containing a liquid as the second consonant (/C+liq/ clusters), it is the second consonant that is omitted, whereas for clusters starting with an /s/ (/sC/ clusters) it is usually the first consonant, that is, the fricative, that is omitted. In the preferential looking experiment the perception of reduced vs. correct /C+liq/ clusters and reduced vs. correct/sC/clusters was therefore compared. Due to the lack of enough prototypical nouns starting with /kn/ this cluster was not included in the experiment. Note that in contrast to previous literature, 'mispronunciations' were created not by substituting one phoneme for another, but by reducing the onset cluster to one of the target consonants. The type of reduction followed the predominant pattern of initial child productions: for /C+liq/-clusters the liquid was omitted while for /sC/-clusters it was the fricative that was omitted (Fikkert, 1994; Jongstra, 2003). We used the PLP to test whether two-year-old Dutch children store onset clusters in their complex form (CCV) even though they produce target words with an onset cluster in a simplified form (CV). The research questions underlying the perception experiment are formulated below.

1. Do two-year-olds perceive the difference between correctly produced and reduced clusters?
2. Is there a difference in the looking behavior towards /C+liq/ cluster trials and /sC/ cluster trials?

To date no experiments have tested the perception of reduced clusters in known words, and only one study has looked at the perception of clusters at all. Archer and Curtin (2011) tested 6- and 9-month-olds' preference for different types of onset clusters in pseudo-words. In the first experiment they contrasted phonotactically well-formed onset clusters like /pl/ and /kl/ with phonotactically ill-formed clusters, like /tl/. The results pointed out that both age groups looked longer at trials containing the ill-formed clusters than at the trials containing only well-formed clusters. Furthermore, only 9-month-olds distinguished between well-formed frequent and infrequent clusters. The authors' conclusion was that while 6-month-olds can discriminate between native and nonnative sound combinations, they are not yet sensitive to type frequency for legal onset clusters. Nevertheless, since this study used only used pseudo-words, the finding that infants are able to distinguish between well- and ill-formed clusters suggest that infants have stored at least some clusters of known words in detail, to allow for generalization to novel words (i.e., pseudo-words). However, to test this assumption it is crucial to test children's perception of known words. This is why we conducted a PLP experiment where the perception of target-like vs. reduced clusters was tested. Furthermore, I did not test random mispronunciations of target clusters, but mispronunciations that are similar to the actual cluster mispronunciations of two-year-olds.

If infants have stored all clusters correctly, they should be sensitive to how clusters are produced and it is therefore expected that they will look longer when correctly produced test items are presented than when test items are presented that are incorrectly produced, with reduced onset clusters. It is possible that infants only notice mispronunciations for one type of onset cluster. Given that /sC/ clusters are generally acquired later than /C+liq/ clusters (Jongstra, 2003), it could be that toddlers only notice mispronunciations for /C+liq/ clusters. However, the opposite is possible too, i.e., toddlers only notice mispronunciations for /sC/ clusters, because the omission of initials consonant, in this case /s/, could be acoustically more salient than omission of a second consonant, as in the case of the omitted liquid. It is, thus, possible that the position of omission plays a role such that omission of a consonant in the first position is more detrimental to word recognition than omission of a consonant in second position. A second possibility why the omission of /s/ is more salient than a liquid is because they differ in acoustic saliency and duration.

We tested two-year olds because at this age they have acquired a vocabulary large enough to contain a variety of words with clusters in the onset. It is also one of the earliest ages at which one can obtain production data in an experimental setting (Hoff et al., 2008). After the perception experiment an elicitation task was carried out to see how the same infants produced words with the two types of clusters. On the basis of their productions the participants were grouped as either "reducers" or "substituters" of consonant clusters. The "substituters" produced consonant clusters, but these clusters were not necessarily segmentally correct. Participants who did not produce any of the words were not grouped.

5.2. Method

5.2.1. Participants

Data from 40 monolingual Dutch children, 20 girls and 20 boys, with a mean age of 24;6 (months; days), ranging between 23;16 and 24;21, were retained for analysis. An additional 18 children were tested but excluded from analysis because they did not complete the test (n = 10) or due to equipment failure (n = 8). All children were reported to have a normal development and were recruited from the Leiden Babylab Database. The study was approved by the Ethical Committee of the faculty of Social Sciences, and parents signed the consent form.

5.2.2. Stimuli

For the perception experiment 27 words were selected: 10 /sC/ - cluster words; 12 /C+liq/-cluster words; 5 filler words that served as distracter words at test. According to Bacchini et al., 2005, most two-year-olds would know these words. See Appendix 1 for a list of the words. For each word, we selected a high-resolution realistic picture with the object appearing on a white background. For the elicitation task we selected six words, of which 3 were /sC/ words and 3 were /C+liq/ words. All of these words were used in the perceptual experiment too.

Auditory stimuli accompanying each picture in the perceptual test were recorded in a soundproof booth, with a sample rate of 44.1 kHz. A female native speaker of Dutch uttered the stimuli in a child-directed manner. All words were recorded in natural carrier-contexts (i.e. not-spliced). Three types of carrier sentences containing the target word were used in the test phase: *Kijk naar de [target], mooi he!*' Look at the [target], isn't it pretty?' or *Zie je een [target]? Vind je het mooi?* 'Do you see a [target]? Do you like it?' and *Kijk, een [target]! Mooi, hè?*, 'Look, a [target]! Isn't it pretty?'. Target words were either a /C+liq/ word like *bril* 'glasses' or *bloem* 'flower' (liquids were always /r/ or /l/), or an /sC/ word like *schoen* 'shoe'. Words were either correctly produced (CC) or reduced (RC). To illustrate, a /C+liq/ word like *bril* was correctly produced as [bʀɪl] and in the incorrect version the cluster was reduced, resulting in [bɪl]; an /sC/ word like *schoen* was correctly produced as [sχun] or incorrectly as [χun]. The mean duration of all correctly pronounced target words was 770 milliseconds (a mean duration of 800 ms for /sC/ words and of 730 ms for /C+liq/ words), while for all mispronounced target words the mean duration was 660 milliseconds (a mean duration of 680 ms for reduced /sC/ words and of 640 ms for reduced /C+liq/ words).

Figure 1: An example from the familiarization phase, with the picture of a shoe (*schoen*).

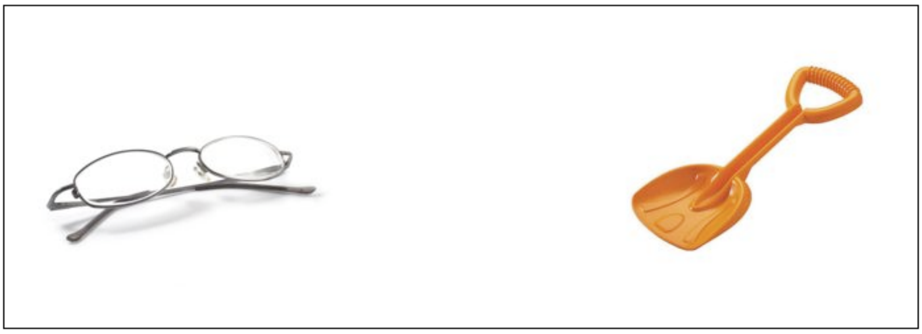


Figure 2: An example from the test phase, where two objects are visible on the screen. In this case the *bril* - *schep* ('glasses' – 'shovel') trial.

5.2.3. Procedure

Children first participated in the perception task, before they were administered the elicitation task. The perception experiment consisted of a familiarization phase, followed by a test phase. The function of the familiarization phase was to make sure that children would recognize the pictures that were presented in the experimental phase. In the familiarization phase, all twenty-seven objects were presented in isolation: visually they were presented with the picture of the target word, slowly moving up and down, while the picture was named - in the correct way, in the recorded female voice. Each familiarization trial lasted for 4 seconds (see Figure 1).

The experimental phase consisted of twenty-five trials (twenty-two test trials; 3 filler trials), in which children saw two objects side by side. The pairs of pictures moved slowly up and down while the auditory stimulus was presented, naming only one of the objects. Each experimental trial lasted for eight seconds. The pairs of pictures were presented two seconds after the beginning of the trial. Paired objects did not overlap in word onset: most pairings comprised objects from two different consonant-cluster pairings (e.g. *bril* - *schep*; 'glasses'-'shovel'; see Figure 2). We controlled for saliency effects; animate objects were paired with other animate objects; inanimate objects with other inanimate objects. Each pairing was presented twice: in total, the participants were invited to look at all objects.

At test, children heard only one pronunciation of each target word: either the correct form or the reduced form. Each child was presented with a mix of correct and reduced test items. Therefore, in order to test all possible trials with correctly produced and reduced clusters two experimental groups were created. Furthermore, to control for the possible diminishing concentration towards the end of the test, each experimental group was tested in two different orders of presentation of the stimuli. There were thus 4 versions of the experiment, and each version was presented to 10 children. In each version, children were presented with a relatively equal amount of correct and reduced pronunciations of target words, for both types of consonant-clusters. In total the experiment counted 4 types of test stimuli, i.e. 4 test conditions: correct C+liquid clusters (CCliq), correct /sC/ clusters, (CCs), reduced C+liquid clusters (RCliq), and reduced /sC/ Clusters (RCs). This is schematized in Table 2 below. All children were presented with all 4 conditions, with 5 or 6 trials per condition. Trials were distributed in a semi-randomized way: two trials from the same condition, or two trials with the same pairing of pictures never immediately followed one another. For a more detailed overview of the different test trials, see Appendices 2 and 3. The familiarization phase lasted for 3 minutes and the experimental phase for 5 minutes.

After participating in the perception experiment, the children were tested with a short word elicitation task. Their utterances were recorded with a Microtrack II digital recorder and an external Microtrack II microphone. First the children were shown pictures of three words, starting either with a /sC/, /Cl/ or a /Cr/ cluster: *stoel* 'chair'; *bloem 'flower',* and *broek* 'trousers'. If they showed interest in participating in the task, they were given the chance to produce 3 more cluster words of the same types. The intention of the word elicitation task was to find out in what stage of development of cluster production the children were, i.e. whether they exhibited cluster reduction in both /sC/ and /C+liq/ clusters, whether they exhibited only /sC/ cluster reduction or whether they mastered the production of both /sC/ and /C+liq/ clusters. Finally, parents filled in a questionnaire concerning their child's receptive and productive vocabulary (short N-CDI's; Zing & Lejaegere, 2003).

Table 2: An overview of all the possible target words used for the picture pair 'bril-schep' (glasses - shovel).

trial type	**cluster type**	**experimental conditions**	**example**
correct cluster	/C+liq/ words	CCliq	bril [bʀɪʟ]
	/sC/ words	CCs	schep [sχɛp]
reduced cluster	/C+liq/ words	RCliq	bil [bɪʟ]
	/sC/ words	RCs	chep [χɛp]

5.2.4. Apparatus

The entire experiment took place in a 2m × 2m soundproof booth. During the experiment, children sat on their caregiver's lap at a distance of 90 cm from the screen. One camera, mounted directly under the screen recorded the children's eye movements. Caregivers wore headphones and listened to a mix of music and backward speech.

An experimenter monitored the session outside the booth while the experiment was run on a Macintosh G4 laptop computer using the Habit X 1.0 software (Cohen et al., 2000). The looking behavior of each participant during the experiment was recorded with a Panasonic camera on a Panasonic DVD recorder. The video recordings of the children’s faces were coded offline by trained scorers using Elan (EUDICO Linguistic Annotator) 3.6.

5.2.5. Scoring

The main interest of the perception experiment was to compare children's looking behavior during the correctly pronounced target words vs. the looking behavior during the reduced target words. Each test trial was divided into two phases: the pre-naming phase measured from the onset of the trial (including the carrier sentence) up to the onset of the target word: 0 – 2,000 ms, and the post-naming phase from 360 ms after the onset of the target word up to 2,000 ms after the onset of the target word: 2,360 – 4,000 ms. The delay of 360 ms after the word onset is considered to be the time that infants need to initiate eye movement in response to speech (Swingley & Aslin, 2000). Only trials during which children fixated both the target and the distractor in the pre-naming phase were taken into account for the final analysis. Two looking-time measures were used: proportion of looking time at the target (PTL) and latency longest look at the target (LLK).

PTL is computed by dividing the total time spent looking at the target by the total time spent looking at either the target or the distractor (Swingley & Aslin, 2000). Latency longest look is the difference between the longest look at the target and the longest look at the distractor (Mani & Plunkett, 2007). The effect of naming on any given trial is the difference in PTL and LLK between the post- and pre-naming phases. A positive difference (post- minus pre-naming phase) indicates that a participant fixated the target relatively more *after* naming than before it was named. In the statistical analysis this difference measure for both the PTL and the LLK measures was used, separately for each of our four conditions (CCs; CCliq; RCs; RCliq). If children would have stored their onset consonant clusters in detail, naming effects are expected to be larger for correctly-produced consonant clusters than for reduced clusters.

For the production analysis children fell into one of 3 categories: (1) Reducers: children who *reduced* clusters; (2) Substituters: children who produced clusters of one or both cluster types, either correctly or with segmental substitutions, and (3) Non-producers: children who refused to participate in the word elicitation task.

5.3. Results: perception of clusters

5.3.1. The Results for PTL measure

A repeated measures (RM) analysis of variance test (ANOVA) was performed, with pronunciation condition (correct cluster vs. reduced cluster) and cluster type (C+liquid cluster vs. /sC/ cluster) as factors.

The analysis of variance revealed that the factor of *cluster type* had a significant effect (F [1,39] = 4.009, p = .05): children looked longer during /sC/ cluster test trials than during /C+liq/ cluster test trials. The factor *pronunciation condition* did not show a significant effect ($p \geq .119$): children did not differentiate between correctly produced or reduced test trials. Neither did the PTL measure point to an interaction between cluster type and pronunciation condition, $p \geq$ .11.

5.3.1.1. Between-subject factors

In the RM ANOVA, where the PTL measure was the dependent variable, the between-subject factors sex and group were also investigated. It is necessary to check that none of the independent subject-related factors are interacting with the dependent factor, in this case the difference PTL measure. There was no interaction between the between-subject factors and the difference PTL measure, for all $p \geq .071$.

5.3.1.2. Planned post-hoc comparisons

One-sample t-tests were performed to examine whether there were naming effects for each cluster-type. Indeed, for all conditions, looking time at the target word in the post-naming phase was significantly longer than chance (i.e. compared to 0), for all $p \leq .034$. This effect was strongest for /sC/ words in the correctly produced condition $t(39) = 4.9$; $p < .001$. In other words, in all the conditions, a naming effect was found and this effect was even stronger in the case of correctly produced /sC/ words.

Naming effects were then compared for each cluster type, either produced correctly or reduced. When words are correctly produced, there is a significantly larger increase for /sC/ words as compared to /C+liq/ words ($t(39) = -2.63$; $p = .012$). When words were produced with reduced clusters, infants' looking behavior was similar ($p \geq .9$). A larger difference PTL measure for correctly produced /sC/ cluster words than for /C+liq/ words means that children looked longer at the target word in the post-naming phase and that this effect was larger for /sC/ words than for /C+liq/ words.

We also wanted to find out whether in the four conditions, children looked longer than chance at the target word in the post-naming phase. This is why the four different conditions were compared to zero by means of one-sample t-tests. The t-tests pointed out that the looking time at the target word in the post-naming phase was significantly longer than chance, for all $p \leq .034$, for the difference PTL measure for /C+liq/ words in the CC and CR condition and for /sC/ words in the CR condition.

Since the ANOVA did not reveal an effect of pronunciation condition for the difference PTL measure, it cannot be concluded that children looked longer in one condition as compared to another condition on the basis of the way the cluster word was pronounced. In Figure 3 the mean values for the difference PTL measure are given, showing how much longer the PTL was in the post-naming phase with respect to the pre-naming phase for words like *bloem* /CCliq/ and *boem /RCliq/*; *stoel* /CCs/ and *toel* /RCs/.

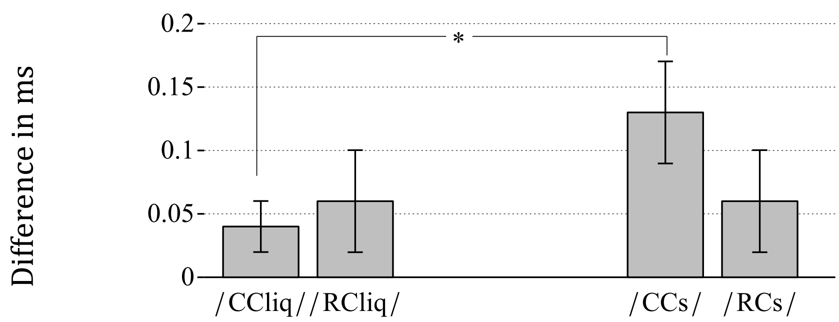


Figure 3: Mean values and SD of the difference PTL score for /C+liq/ words in the correct cluster (CC) vs. /C+liq/ words in the reduced cluster (RC) condition and for /sC/ words in the CC condition vs. /sC/ words in RC condition.

5.3.2. The results for LLK measure

For the LLK measure a repeated measures analysis of variance test (RM ANOVA) was conducted, with naming phase pronunciation condition (correct cluster vs. reduced cluster) and cluster type (liquid cluster vs. /s/ cluster) as independent variables and measure for looking behavior (difference LLK) as dependent variable.

The analysis of variance did not reveal significant effects of *cluster type* and *pronunciation condition*, for all $p \geq .099$. However, there was a significant interaction between the factors *cluster type* and *pronunciation condition* (F [1,39] = 6.51, p = .015). In order to find out in which way this interaction was expressed, a number of paired sample t-tests were carried out, namely for the difference LLK measure for /C+liq/ words in the CC condition vs. the CR condition; for /sC/ words in the CC condition vs. the CR condition; for the /C+liq/ words vs. the /sC/ words in the CC condition, and for the /C+liq/ vs. the /sC/ words in the CR condition. The difference LLK measure turned to be significantly larger for /sC/ words in the CC condition as compared to those in the CR condition $t(39) = 2.28$; $p = .028$; this was not found for /C+liq/ words $p \geq .35$. Furthermore, a significant difference between the /C+liq/ and /sC/ words in the CC condition was found, where the participants looked longer in the post-naming phase at correctly pronounced /sC/ words $t(39) = -3.19$; $p = .003$.

I also wanted to find out whether in the four conditions, the children looked longer than chance at the target word in the post-naming phase. This is why the four different conditions were compared to zero by means of one-sample t-tests. The t-tests pointed out that the looking time at the target word in the post-naming phase was significantly longer than chance, for all $p \leq .012$, difference LLK measure for /C+liq/ words in the CC and CR condition and for /sC/ words in the CR condition. This effect was even stronger for /sC/ words in the CC condition $t(39) = 5.66$; $p = .000$. In other words, in all the conditions, a naming effect was found and this effect was even stronger in the case of correctly produced /sC/ words.

Since the RM ANOVA exhibited an interaction between cluster type and pronunciation condition for the difference LLK measure and the separate t-tests pointed out that this effect was due to participants having their longest look at the correctly pronounced /sC/ target picture in the post-naming phase, as compared to the mispronounced /sC/ target word and as compared to the correctly pronounced /C+liq/ target word. The mean values of the difference LLK measure for target words like *bloem* and *boem*; *stoel* and *toel* are presented in Figure 4 below.

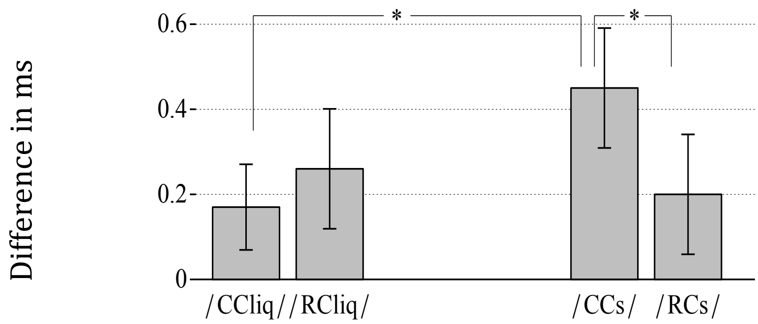


Figure 4: Mean values and SD of the difference LLK score for /C+liq/ words in the correct cluster (CC) vs. /C+liq/ words in the reduced cluster (RC) condition and for /sC/ words in the CC condition vs. /sC/ words in RC condition.

5.3.2.1. Between-subject factors

In the RM ANOVA, where the LLK measure was the dependent variable, the between-subject factors sex and group were also investigated, as it is necessary to check whether any of the independent subject-related factors interact with the dependent factor, in this case the LLK measure. There was no interaction between the between-subject factors and measure of latency longest look, for all $p \geq .115$.

5.3.3. NCDI scores

Parents were asked to fill out two versions of the Dutch-NCDI, NCDI-2A and NCDI-2B, for children between 18- and 24-months. The number of correct scores was then calculated in percentages. For the current correlation test, I only used the results on perception and production from the NCDI-2A version of the test, since the results form NCDI-2A and NCDI-2B were highly similar.

In order to find out if the productive and/or the perceptive vocabulary scores of the two-year-olds were correlated with the difference scores obtained in the perception test, a linear regression test was conducted, with four different measures as independent variables. These were the difference PTL measure for /C+liq/ words in the CC condition and in the RC condition, and the difference measure for /sC/ words in the CC condition and in the CR condition. None of the four cues (difference measure /C+liq/ in CC and in CR condition, difference measure /sC/ words in CC condition and in CR condition) correlated with the vocabulary scores (for all $p > .05$). The same analysis was conducted between the productive NCDI scores and the four PTL difference scores, the perceptive vocabulary and the LLK difference scores, the productive vocabulary and the LLK difference scores and again, no correlation was found (for all $p > .05$).

5.4. Results: Production

Children were grouped based on the way in which they in general produced cluster words in the elicitation task. Eleven children mainly reduced both cluster types (*reducers*), and twelve children produced two consonants in one or both cluster types (*substituters*); seventeen children refused to participate (*non-producers*). Little could be concluded about this last group; it was not the case that they did not talk at all, since their parents reported that they were able to utter words. However, because nothing could be concluded about their cluster production capacities, this group of children was not taken into account in the main analysis below.

Table 3 summarizes findings from the production experiment. Both the reducers and the substituters uttered between 4 and 6 (out of a total of 6, namely 3 /sC/ and 3 /C+liq/ words) words in the elicitation task.

The substituters and the reducers had comparable perceptive and productive vocabulary sizes, as determined by the NCDIs (for both perception and production vocabularies, t-tests turned out to be insignificant, with $p \geq .31$). Non-producers exhibited the lowest perceptive and productive vocabulary scores. Table 3 also shows the production patterns for the substituters and the reducers speaker types.

Table 3: NCDI scores and production patterns for substituters, reducers and non-producers speaker types.

Speaker types	Maximal number of **tokens** that start with a consonant cluster	Number of **children** with production pattern: /C+liq/ correct; /sC/ C1 omitted	Number of **children** with production pattern: /C+liq/ C2 substituted; /sC/ C1 omitted	Number of **children** with production pattern: /C+liq/ C2 omitted; /sC/ C1 substituted	Number of **children** with production pattern: /C+liq/ correct; /sC/ C1 substituted	Number of **children** with production pattern: /C+liq/ C2 omitted; /sC/ C1 or C2 omitted	NCDI-2A prod.	NCDI-2A perc.
Substituter female: 5 male: 7	6	6	3	2	1	0	81 46-100	60 31-93
Reducer female: 8 male: 3	6	0	0	0	0	11	84 50-100	67 26-97

5.5. The link between perception and production

Naming effects were compared for the substituters and the reducers. Since results from the looking behavior with the latency longest look measure suggest that overall, toddlers noticed the mispronunciation for /sC/ words, this measure (i.e. the LLK difference score) was used in the statitical analysis below.

In Table 3 it can be seen that of the twelve cluster reducers, 9 only reduced /sC/ clusters, while they produced target /C+liquid/ clusters correctly - or at least with two consonsants. While these /sC/ cluster reducers did not show different looking times for correct vs. reduced /C+liq/ words ($p \geq .39$), they did look longer at *correctly* pronounced /sC/-words than at reduced /sC/-words ($t(11) = 5.97$, $p = .0001$). In other words, the children who tended to reduce /sC/ clusters were sensitive to the difference between correctly produced and reduced /sC/ words, but not to the difference between correctly produced and reduced /C+liq/ words.

Similar paired-sample t-tests were carried out for the substituters. Here we did not find significant differences, neither for the /C+liq/ words, nor for the /sC/ words, (for both $p \geq .64$). Although this group of children has reached a higher level in their production performance than the reducers, they did not look significantly longer to correctly produced words: they increased their looks only marginally. However, this increase is not significantly different from the reducers' increase for the correctly produced /sC/-clusters ($p \geq .44$). This suggests that the substituters reveal the same pattern as reducers, but to a lesser extent. I also used unpaired t-tests to test whether substituters and reducers exhibited a naming effect for either /C+liq/ or /sC/ words. I found that both reducers and substituters showed a naming effect for /sC/ words ($t(11) = -3.44$, $p = .006$ and $t(12) = -3.19$, $p = .008$, respectively) but not for /C+liq/ words (for both production types $p \geq .26$).

To sum up, I found an asymmetry between production patterns and perception patterns, where children who reduced /sC/ clusters in their productions showed awareness of the difference between correctly and incorrectly pronounced /sC/ clusters, while children who were more advanced in their realization of clusters did not. With respect to /C+liq/ words, no significant differences were found.

5.6. Discussion

In this chapter we looked at children's looking behavior for correctly produced and mispronounced /sC/ words and /C+liq/ words in a preferential looking paradigm experiment. In the data we obtained, it was observed that although all four conditions elicited naming effects (i.e. increase at looking at target), there was a significant difference between cluster types when measuring children’s longest looks. The largest naming effect was obtained for correctly produced /sC/ words and it differed significantly from the effect for the incorrect pronunciation of /sC/ words.

Recall that in other PLP studies, longer fixation at the target is found when it is named correctly as opposed to when the target is mispronounced (Swingley & Aslin, 2000) and that bigger violations in the target word lead to shorter fixation at the target picture (White & Morgan, 2008). In studies that found that infants are sensitive to mispronunciations, it is often concluded that they have a detailed lexical representation of the correct target word. Thus, according to these studies, longer fixation at a target picture named correctly (as opposed to fixation at the target named with a mispronounced label) points to a detailed lexical representation of the target word. If we apply this line of reasoning to the difference in looking times obtained for /sC/ words in the present study, we would have to conclude that 24-month-old Dutch children have a detailed representation of the onset cluster in the /sC/ words, since they are sensitive to C_1 deletion.

However, if we use this interpretation for /sC/ clusters, then the question arises why the same subjects did not exhibit a significant difference in looking time for correctly produced /C+liquid/ words vs. mispronounced /C+liquid/ words, with longer looks for correct productions, while most of these subjects even produced two consonants when attempting these target clusters. The 24-month-olds looked equally long at the picture of a 'bread' when they were presented with /bʀo:t/ as when they were presented with /bo:t/, suggesting that they found the word label lacking the C_2 in the onset a "good enough" exemplar of the Dutch word *brood*. The apparent insensitivity to an incorrect form of *brood* vs. the sensitivity to a correct form of *stoel* could be explained by saliency. It could be that children are only sensitive to the correctness of consonants at word edges (Swingley & Aslin, 2000; Swingley, 2005) and of the vocalic nucleus (Mani & Plunkett, 2007), and that they are less sensitive to consonants in non-edge positions, like the C_2 in the present study. The difference in sensitivity could also be enhanced by the higher acoustic saliency of the fricative in a C_1 position.[2]

While one of these types of saliency could be involved in the results here, it is remarkable, in Figure 4, that the LLK score for correct /C+liq/ words is not only comparable to the LLK score for mispronounced /C+liq/ words, but also to the score for incorrect /sC/ words. This suggests that listeners are equally "uninterested", in the *correct* naming of *brood* as in the *incorrect* naming of *stoel*. The difference between correctly produced *brood* and correctly produced *stoel* for the LLK measure is a striking 230 ms. This difference is not caused by a general preference for /sC/ words (or for the objects which picture /sC/ words) as the looking times to objects in the two conditions did not significantly differ from one another in the pre-phase ($p \geq .36$). In the post-phase, however, the LLK measure was significantly higher for correctly produced /sC/ words than for correctly produced /C+liq/ words ($t(39) = -3.03$, $p = .004$). This difference cannot be explained by the mean duration of /sC/ words either; these words were, on the whole, 55 ms longer than /Cliq/ words, but this is insufficient to explain the longer LLK of 230 ms for correctly produced /sC/ from the LLK for correctly produced /C+liq/ targets.

The short looking times at correct /C+liq/ words suggest an alternative hypothesis, according to which the results in this chapter point to a learning effect for /sC/ clusters. These clusters are usually acquired later than /C+liq/ clusters (Fikkert, 1994). The learning effect results from the learners' comparison of their own reduced forms of target /sC/ words, like [tul] for target *stoel* (chair) /stul/, to the perceived correct form that they are presented with, [stul]. If their own segmental representation of the word does not contain /s/, and they notice the mismatch of their own form compared to the form they

---

[2] According to the sonority scale (Selkirk, 1984) /s/ is less sonorous than liquids. However, when /s/ is in word initial position, it is acoustically more salient than the liquid following a plosive in /C+liq/ clusters.

are presented with, this might lead to increased attention for this correct form, because they are updating their segmental representation, and therefore to longer looking times in this condition. With respect to /C+liq/ words, we would then hypothesize that the children already have segmental representations that include the complete cluster. The correct forms therefore do not invite an update of these representations. However, the reduced forms, like [bot] for *brood* (bread) /brot/, are “good enough” perceptual exemplars of their segmental representations of these forms. The reasons for this apparent indifference to correct versus incorrect forms in the case of /C+liquid/ forms could again be due to the fact that the C_2 position is less salient. Using a more sensitive experimental technique, like EEG, could show that children are, in fact, sensitive to the missing liquid in the reduced test items.

Furthermore, it is interesting to note that production data of the same participants in our experiment show that the strongest naming effect for correct /sC/ clusters is found for the eleven speakers (out of the total of twenty-three who were willing to participate) who reduced /sC/ clusters in their own productions. The remaining twelve speakers who tended to exhibit substitution in their cluster realizations also show the naming effect for correct /sC/ words but to a much lesser extent. The fact that the /sC/ naming effect was stronger among the reducers, is in line with the learning effect hypothesis.

5.7. Conclusion

In this chapter I studied the way in which 24-month-olds perceive different reduced onset clusters. Furthermore, I investigated whether they react differently to /C+liq/ onset cluster reductions, on the one hand, and reductions of /sC/ onset clusters, on the other hand. The results from this experiment suggest that the participants are able to differentiate between correctly produced and reduced /sC/ cluster words but not between correctly produced and reduced /C+liq/ words. Interestingly, it was mainly the group of children who reduced target /sC/ clusters in their own productions that showed the highly increased looking times for correctly produced /sC/ test items.

Four accounts for these data were considered. The first account was based on the assumption in the literature on similar preferential looking experiments, that longer looking times for correct pronunciations test stimuli indicate detailed segmental representations of those stimuli. This would indicate that the longer looking times that were found for the correct production of /sC/ test stimuli point to the children in the study having a detailed lexical representation of /sC/ words. This account does not seem to be supported by the production data of the children in the study, which showed a general pattern of fairly correct productions of target /C+liquid/ clusters, next to mostly reduced productions of target /sC/ clusters. The second and third accounts concern the lack of sensitivity to the C_2 omission in target /C+liq/ stimuli. The second account states that the C_2 position is less salient, and therefore is of less importance for young children's word recognition. The third account states that the sensitivity to the presence of [s] in the correctly produced /sC/ words is specifically due to the high acoustic saliency of the segment /s/ in word initial position. While both saliency accounts could explain the perception data, they again do not seem to fully comply with the production data, unless the omission of /s/ in the productions of target /sC/ clusters of most of the participating children is mainly due to speech production constraints. The saliency hypothesis could be tested in an experiment with the same stimuli but testing the reverse reduction patterns. Thus, in *stoel* C_2 is omitted and becomes [sul] and in *brood* C_1 is omitted and becomes [ʀoːt]. Although these simplification patterns are uncommon in Dutch children's productions, it would be interesting to see whether we would now find the opposite pattern in looking times, with the longest looking times for the reduced /C+liq/ stimuli, that now lack their C_1. If this would be the case, the specific saliency account, stating that /s/ in C_1 position is highly salient and attracts the attention when it is present, would be automatically rejected. Finally, the fourth account suggests that the longer looking times for correctly produced /sC/ test stimuli result from online learning. The idea behind this hypothesis is that the segmental representation of target /sC/ clusters for children who reduce these clusters to [C] in their productions is the reduced form /C/. The longer looking times for correct /sC/ test items can then be taken to indicate that they are paying attention to the mismatch between their own [C] production and the perceived [sC] production of the word referring to the target picture, and that they might even be updating their segmental representation. The online awareness of a mismatching representation that needs to be updated when being presented with correct /sC/ test trials, then, causes the difference between the looking times for correct /sC/ and /C+liq/ forms. An additional assumption is that the children in this experiment are already aware of the complex onset of /C+liq/ clusters, as evidenced by their productions of words starting with a /C+liquid/ cluster, which usually contains a consonant cluster.

In order to test the learning hypothesis, the experiment from this chapter should be performed with a younger group children, who are still reducing target /C+liquid/ clusters in their production. The expectation is that they would exhibit the same effect for the correctly produced /C+liq/ test stimuli that was found here for the correctly produced /sC/ test stimuli.

Appendix 1: A list of the 27 words used in the familiarization phase, which were presented in random order.

Test words used in this experiment		
/Cr-/ and /Cl-/ words	**/sC-/ words**	**distractor words**
vlag	schep	baby
brood	schaar	bal
broek	schoen	auto
trui	speen	poes
bril	step	mier
blok	stoel	
trap	spin	
klok	slang	
kraan	schaap	
trein	slak	
fles		
bloem		

Appendix 2: A list of the 25 trials that were presented in the experimental phase, for the 1st experimental group. The table gives information about the picture pairs that were presented in each trial, the auditory stimulus (annotated using Dutch orthography) that was part of the trial and the condition of each trial, where CP stands for correct pronunciation and MP for mispronunciation, 'liq' refers to /C+liq/ words and 's' to /Cs/ words. The trial order of the table is identical to the 1st experimental order used during the experimental phase.

trial number	left picture	right picture	auditory stimulus	condition
1	bloem	step	step	CPs
2	vlag	broek	vlag	CPliq
3	brood	flesje	bood	MPliq
4	poes	schaap	schaap	CPs
5	spin	slang	lang	MPs
6	bril	schep	schep	CPs
7	auto	baby	baby	CPcontrol
8	broek	vlag	boek	MPs
9	trui	speen	trui	CPliq
10	schoen	kraan	choen	MPs
11	slang	spin	spin	CPs
12	schep	bril	bil	MPliq
13	blok	schaar	bok	MPliq
14	stoel	trap	trap	CPliq
15	baby	auto	paku	MPcontrol
16	klok	trein	klok	CPliq
17	bal	poes	bal	CPcontrol
18	trap	stoel	toel	MPs
19	kraan	schoen	kraan	CPliq

20	mier	slak	slak	CPs
21	trein	klok	tein	MPliq
22	speen	trui	peen	MPs
23	flesje	brood	fles	CPliq
24	step	bloem	boem	MPliq
25	schaar	blok	schaar	CPs

Appendix 3: A list of the 25 trials that were presented in the experimental phase, for the 2nd experimental group. The table gives information about the picture pairs that were presented in each trial, the auditory stimulus (annotated using Dutch orthography) that was part of the trial and the condition of each trial, where CP stands for correct pronunciation and MP for mispronunciation, 'liq' refers to /C+liq/words and 's' to /Cs/ words. The trial order of the table is identical to the 2nd experimental order used during the experimental phase.

trial number	left picture	right picture	auditory stimulus	condition
1	blok	schaar	chaar	MPs
2	spin	slang	slang	CPs
3	stoel	trap	stoel	CPs
4	slak	mier	lak	MPs
5	vlag	broek	broek	CPliq
6	trap	stoel	tap	MPliq
7	schoen	kraan	kaan	MPliq
8	trein	klok	kok	MPliq
9	brood	flesje	fes	MPliq
10	auto	baby	auto	CPcontrol
11	bloem	step	bloem	CPliq
12	bril	schep	bril	CPs
13	trui	speen	tui	MPliq
14	schep	bril	chep	MPs
15	klok	trein	trein	CPliq
16	step	bloem	tep	MPs
17	kraan	schoen	schoen	CPs
18	poes	bal	diem	MPcontrol
19	flesje	brood	brood	CPliq

20	baby	auto	zoma	MPcontrol
21	schaap	poes	chaap	MPs
22	speen	trui	speen	CPs
23	broek	vlag	vag	MPliq
24	schaar	blok	blok	CPliq
25	slang	spin	pin	MPs

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.

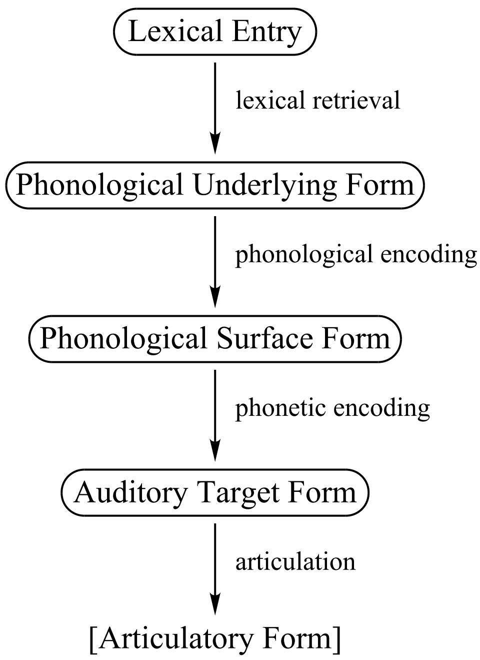


Figure 1: The speech production model, based on Boersma and Hamann (2009), Boersma (2011) and Levelt et al. (1999).

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 [sʷxʷɔmˑəɫ].

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 [sʷxʷɔmˑəɫ] is mapped onto the representation [ˈsxɔməɫ] 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.

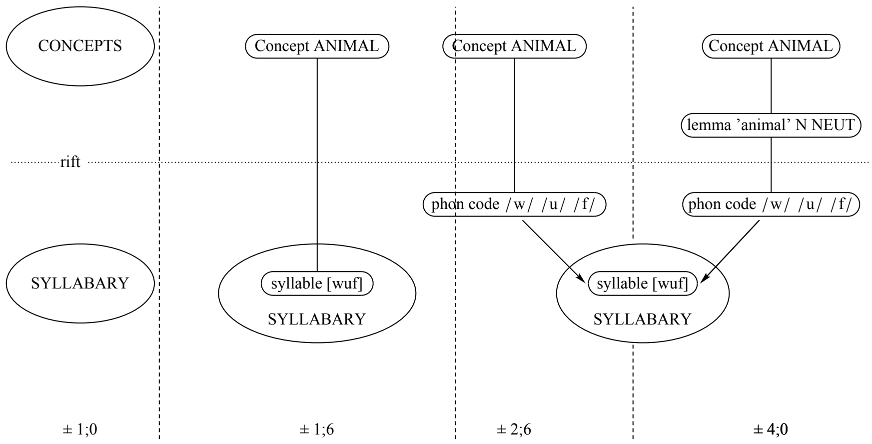


Figure 2: Four-staged model as suggested by Levelt (1998), visualizing how the child speech production model is built up.

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:

↗CCV(C) → (C)VCC ↘
CV → CVC → V CCVCC
↘ (C)VCC → CCV(C) ↗

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.

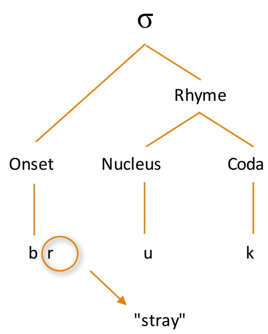
6.5 The developing speech production mechanism

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

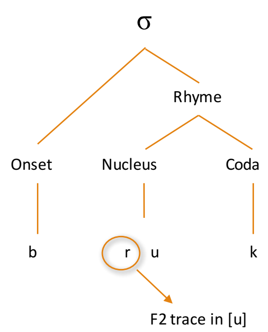


Stage: /r/ = F2 trace

Phonological encoding: 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.

Phonetic encoding: 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.

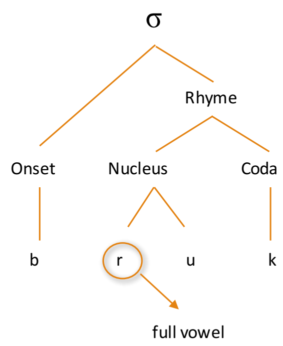


Stage: /r/ = full vowel

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



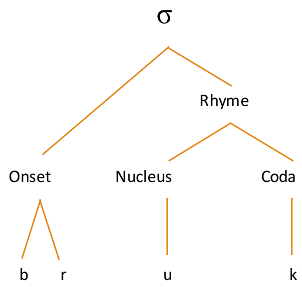
Stage: epenthesis in target cluster

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



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.

References

Altvater-Mackensen, N. & Mani, N. (2013). Word-form familiarity bootstraps infant speech segmentation. *Developmental Science, 16, 6*, 980-990.

Adriaans, F. W. (2011). *The induction of phonotactics for speech segmentation. Converging evidence from computational and human learners*. PhD thesis. LOT: Utrecht.

Archer, S. L. & Curtin, S. (2011). Perceiving onset clusters in infancy. *Infant behavior and development, 34*, 534-540.

Bacchini, S., Boland, T., Hulsbeek, M., Pot, H. & Smits, M. (2005). *Duizend-en-een woorden, de allereerste Nederlandse woorden voor anderstalige peuters en kleuters. Stichting leerplanontwikkeling. Een gefundeerde woordenlijst gefuseerd naar verwerving*. [online document]. http://www.slo.nl/downloads/archief/Duizend-en-een-woorden.pdf/

Ball, M. J., Lowry, O. & McInnis, L. (2006). Distribution and stylistic variation in /r/-misarticulations: a case study. *Clinical Linguistics & Phonetics, 20, 2/3*. 119-124.

Bailey, T.M. & Plunkett, K. (2002). Phonological specificity in early words. *Cognitive Development, 17,* 1265-1282.

Berko, J. & Brown, R. (1960). Psycholinguistic research methods. In P. H. Mussen (ed.) *Handbook of research methods in child development*. New York: Wiley. 517–557.

Blevins, J. (1995). The syllable in phonological theory. In Goldsmith J (ed.) *The handbook of phonological theory*. Oxford: Blackwell. 206–244.

Boersma, P. (1998). *Functional phonology: formalizing the interactions between articulatory and perceptual drives*. PhD thesis. University of Amsterdam.

Boersma, P. (2011). A programme for bidirectional phonology and phonetics and their acquisition and evolution. In A. Benz & J. Mattausch (eds.) *Bidirectional Optimality Theory*. Amsterdam: John Benjamins. 33–72.

Boersma, P. & Weenink, D. (2008). Praat: Doing phonetics by computer (Version 5.0.10) [Computer program]. http://www.praat.org.

Boersma, P. & Hamann, S. (2009). Loanword adaptation as first language phonological perception. In A. Calabrese & W. L. Wetzels (eds.) *Loanword phonology*, 11-58. Amsterdam: John Benjamins.

Bowey, J. A. (2001). Nonword repetition and young children's receptive vocabulary: A longitudinal study. *Applied Psycholinguistics*, *22*(03), 441-469.

Browman C. P. & Goldstein, L. (1989). *Articulatory gestures as phonological units*. Haskins Laboratories Status Report on Speech Research, SR-99, 100, 69-101.

Browman C. P. & Goldstein, L. (1992). Articulatory phonology: An overview. *Haskins Laboratories Status Report on Speech Research, 111-112*, 23-42.

Carter, A. & Gerken, L. (2004). Do children's omissions leave traces? *Journal of Child Language, 31*, 561–586.

Catts, H. W. & Kamhi, A. G. (1984). Simplification of /s/ + stop consonant clusters: a developmental perspective. *Journal of Speech, Language, and Hearing Research, 27*, 556-561.

Chaney, C. (1988). Acoustic analysis of correct and misarticulated semivowels. *Journal of Speech, Language, and Hearing Research, 31*, 275-287.

CHILDES (Child Language Data Exchange System). URL: http://childes.psy.cmu.edu/
Cohen, L. B., Atkinson, D. J., Chaput, H. H., (2000). Habit X: A new program for obtaining and organizing data in infant perception and cognition studies (Version 1.0)[Computer software] Austin: University of Texas

De referentie van Chomsky is: Chomsky, N. (1981). *Lectures on Government and Binding*. Mouton de Gruyter.

Cooper, F. S., Liberman, A. M. & Borst, J. M. (1951). The interconversion of audible and visible patterns as a basis for research in the perception of speech. *Proceedings of the National Academy of Sciences, 37*, 318-325.

Davis, B. L. & MacNeilagde, P. F. (1990). Acquisition of correct vowel production: a quantitative case study. *Journal of Speech and Hearing Research, 33*, 16-27.

Davis, B. L. & MacNeilage, P. F. (1995). The articulatory basis of babbling. *Journal of Speech and Hearing Research, 33,* 1199-1211.

Dell, G. S., Martin, N. & Schwartz, M. F. (2007). A case-series test of interactive two-step model of lexical access: Predicting word repetition from picture naming. *Memory and Language, 56*, 490-520.

Demuth, K. & Fee, J. (1995). Minimal prosodic words in early phonological development. Brown University, Providence, Rhode Island and Dalhousie University, Halifax, Nova Scotia.

Demuth, K. (1996) The Prosodic Structure of Early Words. In J. Morgan & K. Demuth (eds.). *From Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*. Hillsdale, N.J.: Lawrence Erlbaum Associates. 171-184.

Den Ouden, D.B. (2002). *Phonology in Aphasia; Syllables and Segments in Level-specific Deficits*. Doctoral Dissertation, GRODIL 39, University of Groningen.

Dresher, B. E., & Kaye, J. D. (1990). A computational learning model for metrical phonology. *Cognition*, *34*, *2*, 137-195.

Eilers, R. E. & Minifie, F. D. (1975). Fricative discrimination in early infancy. *Journal of Speech, Language and Hearing Research, 18*, 158-167.

Ellis, A. W. & Young, A. W. (1988). *Human cognitive neuropsychology*. London: Taylor & Francis.

Fennell, C. T. & Werker, J. F. (2003). Early word learners' ability to access phonetic detail in well-known words. *Language and Speech, 46*, 245–264.

Ferguson, C. A. & Macken, M. A. (1983). The role of play in phonological development. In K. E. Nelson (ed.) *Children's language*. Hillsdale NJ: Erlbaum.

Fikkert, P. (1994). *On the acquisition of prosodic structure.* PhD thesis. The Hague: Holland Academic Graphics.

Gafos, A. (2002). A grammar of gestural coordination. *Natural Language & Linguistic Theory* 20, 269–337.

Gay, T., Lindblom, J. & Lubker, J. (1981). Production of bite-block vowels : Acoustic equivalence by selective compensation. *Journal of the Acoustic Society of America, 69*, 802–810.

Gerrits, E. (2010). Acquisition of /s/-clusters in Dutch speaking children with phonological disorders. Clinical Linguistics and Phonetics, 24, 3. 199-209.

Gilbers, D. G. & Den Ouden, D. B. (1994). Compensatory lengthening and cluster reduction in first language acquisition: A comparison of different analyses. In A. de Boer, H. de Hoop, and H. de Swart (eds.) *Language and Cognition 4*, *Yearbook 1994 of the research group for Linguistic Theory and Knowledge Representation of the University of Groningen*. 69–82.

Gnanadesikan, R., Kettenring, J. R. and Tsao, S. L. (1995). Weighting and selection of variables for cluster analysis. *Journal of Classification 12, 1,* 113-136.

Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M. & Gordon, L. (1987). The eyes have it. Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language, 14*, 23-45.

Golinkoff, R. M., Ma, W., Song, L. & Hirsh-Pasek, K. (2013). Twenty-five years using the intermodal preferential looking paradigm to study language acquisition: What have we learned? *Perspectives on psychological science, 8, 3,* 316-339.

Greenlee, M. (1974). Interacting processes in the child's acquisition of stop-liquid clusters. *Papers and Reports on child language development, No. 7*.

Gulian, M. (in preparation). Dutch /r/ in word onset clusters: Acoustic analyses of [r] and [ʀ] produced by female speakers.

Gulian, M. & Levelt, C. C. (2009). An acoustic analysis of child language productions with reduced clusters. In J. Chandlee, M. Franchini, S. Lord, and G.-M. Rheiner (eds.) *Boston University Conference on Language Development 33: on-line proceedings supplement.* [http://www.bu.edu/bucld/proceedings/supplement/vol33/]

Gulian, M. & Levelt, C. C. (2011). Temporal measures of reduced /sC/-clusters in toddler speech: Evidence for a detailed lexical specification. *17th International Congress of Phonetic Sciences, Hong Kong*, 787–790. [http://www.icphs2011.hk/ICPHS_CongressProceedings.htm]

Gulian, M., Junge, C. & Levelt, C. C. (2014). Two-year-olds distinguish snakes from nakes but not trains from tains. *Proceedings of the 38th Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.

Harrington, J. (2010). Acoustic phonetics. In W. D. Hardcastle, J. Laver, and F. E. Gibbon (eds.) *Handbook of phonetic sciences: Second edition*. Chichester: Blackwell Publishing. 81–130.

Havy, M. & Nazy, T. (2009). Better processing of consonantal over vocalic information in word learning at 16 months of age. *Infancy, 14, 4.* 439-456.

Hayward, K. (2000). *Experimental phonetics.* Pearson Education Limited: Essex.

Hoff, E., Core, C. & Bridges, K. (2008). Non-word repetition assesses phonological memory and is related to vocabulary development in 20 and 24-month-olds. *Journal of Child Language, 35*, 903-916.

Jongstra, W. (2003). Variable and stable clusters: Variation in the realization of consonant clusters. *Canadian Journal of Linguistics, 48, 3/4*, 265-288.

Junge, C. (2011). *The relevance of early word recognition: Insights from the infant brain*. PhD thesis. Enschede: Ipskam drukkers

Kager, R., & Zonneveld, W. (1986). Schwa, syllables, and extrametricality in Dutch. *The Linguistic Review*, *5*, *3*, 197-221.

Kent, R. D. & Read, C. (2002). *The acoustic analysis of speech.* Albany: Thomson Learning.

Knight, R. A., Villafaña-Dalcher, C. V. & Jones, M. J. (2007). A real-time case study of rhotic acquisition in Southern British English. *16th International Congress of Phonetic Sciences, Saarbrücken*, 1581-1584.

Kohn, S. E., & Goodglass, H. (1985). Picture-naming in aphasia. *Brain and language*, *24*, *2*, 266-283.

Ladefoged, P. (2001). *Vowels and Consonants: An introduction into the sounds of languages.* Backwell: Los Angeles

Ladefoged, P. & Maddieson, I. (1996). The sounds of the worlds' languages. Cambridge, MA, and Oxford: Blackwell.

Lawson, E., Stuart-Smith, J. & Scobbie, J. (2008). Articulatory insights into language variation and change: preliminary findings of derhoticization in Scottish English. *University of Pennsylvania Working Papers in Linguistics, 14, 2*.

Levelt, C. C. (1994). *On the acquisition of place.* PhD thesis. The Hague: Holland Academic Graphics.

Levelt, C. C., Schiller, N. O. & Levelt. W. J. (2000). The acquisition of syllable types. *Language acquisition 8. 3*, 237-264.

Levelt, C. C. (2012). Perception mirrors production in 14- and 18-month-olds: The case of coda consonants. *Cognition, 123*, 174-179.

Levelt, W. J. M. (1998). The genetic perspective in psycholinguistics or where do spoken words come from? *Journal of Psycholinguistics Research, 27, 2*, 167-180.

Levelt, W. J. M., Roelofs, A. & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22, 1*, 1-75.

Levelt, C. C. Schiller, N. O. & Levelt, W. J. M. (1999), A developmental grammar of syllable structure in the production of child language. *Brain and Language, 68*, 291-299.

Lindau, M. (1985), The story of /r/. In V. A. Fromkin (ed.) *Phonetic linguistics. Essays in honor of Peter Ladefoged*. Orlando: Academic Press.

Lleo, C. & Prinz, M. (1996). Consonant clusters in child phonology and the directionality of syllable structure assignment. *Journal of Child Language, 23, 01,* 31-56.

Lowenstein, J. H. & Nittrouer, S. (2008). Patterns of acquisition of native voice onset time in English-learning children. *Journal of the Acoustic Society of America, 124,* 1180-1191.

Macken, M. A. & Barton, D. (1980). The acquisition of the voicing in English: A study of voice onset time in word-initial stop consonants. *Journal of Child Language, 7*, 41–74.

MacNeilage, P. (1981). Feedback in speech production: An ecological perspective. In: T. Myers, J. Laver & J. Anderson (eds.) *The cognitive representation of speech*. Amsterdam: North-Holland.

Mani, N., Mills, D. L. & Plunkett, K. (2012). Vowels in early words: an event-related potential study. *Developmental Science, 15,* 2-11.

Mani, N. & Plunkett, K. (2007). Phonological specificity of vowels and consonants in early lexical representations. *Journal of Memory and Language, 57,* 252-272.

McLeod, S., Van Doorn, J. & Reed, V. A. (1998). Homonyms in children's productions of consonant clusters. In W. Ziegler & K. Deger (eds.) *Clinical phonetics and linguistics*. London: Whurr. 106–114.

McLeod, S., Van Doorn, J. & Reed, V. A. (2001). Normal acquisition of consonant clusters. *Language pathology, 10, 2*, 99-110.

Menn, L. (1976). *Pattern, control, and contrast in beginning speech: A case study in the development of word form and word function*. Doctoral Dissertation, University of Illinois, Urbana. Published, Bloomington: Indiana University Linguistic Club (1979).

Metsala, J. L. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Educational Psychology*, *91*, *1*, 3.

Mills, D. L., Prat, C., Zangl, R., Stager, C. L., Neville, H. J., & Werker, J. F. (2004). Language experience and the organization of brain activity to phonetically similar words: ERP evidence from 14-and 20-month-olds. *Journal of Cognitive Neuroscience*, *16*, 1452-1464.

Munson, B., Edwards, J. & Beckman, M. E. (2005). Relationships between nonword repetition accuracy and other measures of linguistic development in children with phonological disorders. *Journal of Speech, Language and Hearing Research, 48*, 61-78.

Newman, R. S. (2008). The level of detail in infants' word learning. *Current Directions in Psychological Science, 17*, 229–232.

Nijland, L. & Maasen, B. (2005). Syllable planning motor programming deficits in developmental apraxia of speech. In R. J. Hartsuiker, R. Bastiaanse, A. Postma, F. Wijnen (eds.), *Phonological encoding and monitoring in normal and pathological speech*. Psychology Press: New York.

Oller, D. K. & MacNeilage, P. (1983). Development of speech production: perspectives from natural and perturbed speech. In: P. F. MacNeilage (ed.), *The production of speech*. New York: Springer. 91–108.
Ouden, D. B., Den, (2002), Phonology in aphasia: syllables and segments in level specific deficits. PhD thesis. Enschede: Ipskamp.

Paradis, J. (2011). Individual differences in child English second language acquisition: Comparing child-Internal and child-external factors. *Linguistic Approaches to Bilingualism. 1, 3*, 213-237.

Pater, J. & Barlow, J. (2003). Constraint conflict in cluster reduction. *Journal of Child Language, 30*, 487–526.

Plug, L. & Ogden, R. (2003). A parametric approach to the phonetics of postvocalic /r/ in Dutch. *Phonetica, 60*, 159–186.

Prince, A., & Smolensky, P. (1991). *Connectionism and Harmony Theory in linguistics*. University of Colorado, Boulder, Department of Computer Science.

Pruthi, T., Epsy-Wilson, C.Y. & Story, B. H. (2007). Simulations and analysis of nasalized vowels based on magnetic resonance imaging data. *Journal of the Acoustic Society of America, 121*, 3858–3873.

Richtsmeier, P. T. (2010). Child phoneme errors are not substitutions. Toronto Working Papers in Linguistics, 33: [http://twpl.library.utoronto.ca/index.php/twpl/article/view/6889]

Rong, P. & Kuehn, D. P. (2010). The effect of oral articulation on the acoustic characteristics of nasalized vowels. *Journal of the Acoustic Society of America, 127, 4*, 2543-2553.

Rose, Y., MacWhhinney, B., Byrne, R., Hedlung, G., Maddocks, K., Obrien, P., & Warenham, T. (2006). Introducing Phon: A software solution for the phonological acquisition. In D. Bamman, T. Magnitskaya, & C. Zaller (eds.). *Proceedings of the 30th Annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.

Roy, P. & Chiat, S. (2004). A prosodically controlled word and nonword repetition task for 2- and 4-year-olds: Evidence from typically developing children. *Journal of Speech, Language and Hearing Research, 47*, 223-234.

Scobbie, J. M. (1998). Interactions between the acquisition of phonetics and phonology. In C. Gruber et al. (eds.) *CLS 34: The panels*, 343–358.

Scobbie, J. M. & Sebregts, K. (2011). Acoustic and phonological perspectives of rhoticity and /r/ in Dutch. In R. Folli and C. Ulbrich (eds.) *Interfaces in linguistics: New research perspectives*. Oxford: Oxford University Press. 257–277.

Sebregts, K. (2015). *The sociophonetics and phonology of Dutch r*. PhD thesis. Utrecht: LOT.

Selkirk, E. (1984). On the major class features and syllable theory. In M. Aronoff and R. Oehrle (eds.) *Language sound structure.* Cambridge MA, London: MIT Press. 107–136.

Shriberg, L. D., & Kwiatkowski, J. (1982). Phonological disorders IIIA procedure for assessing severity of involvement. *Journal of Speech and Hearing Disorders*, *47*, *3*, 256-270.

Shriberg, L. D. (1993). Four new speech and prosody-voice measures for genetics research and other studies in developmental phonological disorders. *Journal of Speech, Language, and Hearing Research*, *36*, *1*, 105-140.

Shriberg, L. D., Austin, D., Lewis, B. A., McSweeny, J. L., & Wilson, D. L. (1997). The Percentage of Consonants Correct (PCC) MetricExtensions and Reliability Data. *Journal of Speech, Language, and Hearing Research*, *40*, *4*, 708-722.

Smit, A. B. (1993). Phonological error distributions in the Iowa-Nebraska articulation norms: Word-initial consonant cluster. *Journal of Speech, Language and Hearing Research, 36, 5*, 931-947.

Smolensky, P. (1996). On the comprehension/production dilemma in child language. *Linguistic Inquiry, 27*, 720–731.

Song, J. Y. & Demuth, K. (2008). Compensatory vowel lengthening for omitted coda consonants: A phonetic investigation of children's early representations of prosodic words. *Language and Speech, 51*, 385–402.

Stackhouse, J. & Wells, B. (1997). *Children speech and literacy difficulties: A psycholinguistic framework.* London: Whurr.

Stager, C. L. & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature, 388*, 381-382.

Steele, J. (2002). L2 learners' modification of target language syllable structure: prosodic licensing effects in interlanguage phonology. In A. James & L. Leather (eds.) *New Sounds 2000: Proceedings of the 4th International Symposium on the Acquisition of Second Language Speech*. Klagenfurt: University of Klagenfurt. 315-324.

Studdert-Kennedy, M. (1987). The phoneme as a perceptuomotor structure. In A. Allport, D. MacKay, W. Prizz, W & E. Scheerer (eds.) *Language perception and production*. Academic Press: London.

Swingley, D. & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children, *Cognition, 76*, 147–166.

Swingley, D. & Aslin, R. N. (2007). Lexical competition in young children's word learning. *Cognitive Psychology, 54*, 99–132.

Swingley, D. (2003). Phonetic detail in the developing lexicon. *Language and Speech, 46*, 265–294.

Swingley, D. (2005). 11-month-olds' knowledge of how familiar words sound. *Developmental science*, *8*, *5*, 432-443.

Swingley, D. (2009). Onset and codas in 1.5-year-olds' word recognition. *Journal of memory and language, 60*, 252-269.

Trehub, S. E. & Shenfield, T. (2007). Acquisition of early words from single-word and sentential contexts. *Developmental Science, 10*, *2*, 190-198.

Trommelen, M. (1984). *The Syllable in Dutch*. Dordrecht: Foris

Van der Feest, S. V. H. (2007). *Building a phonological lexicon: the acquisition of the Dutch voicing contrast in perception and production*. PhD thesis. Radboud University Nijmegen.

Van der Stelt, J. M., Zajdo, K., Wempe, T. G. & Pols, L. C. W. (2005). Exploring the acoustic vowel space in two-year-old children. Results from Dutch and Hungarian. *Speech Communication, 47*, 143–159.

Van Leeussen, J. W. (2009). *Investigating consonant inventories with acoustic and articulatory models*. MA thesis. University of Amsterdam.

Vance, M., Stackhouse, J. & Wells, B. (2005). Speech-production skills in children ages 3-7 years. *International Journal of Language and Communication Disorders, 40, 1*, 29-48.

Wanrooij, K., Boersma, P. & van Zuijen, T. L. (2014). Fast phonetic learning occurs already in two-to-three month olds: an ERP study. *Frontiers in psychology, 5, 77*, 1-12.

Weenink, D. J. M., (1985). Formant analysis of Dutch vowels from 10 children. *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam*, 9, 45–52.

Werker, J. F. & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant behavior and Development, 7*, 49-63.

White, K. S. & Morgan, J. L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language, 59*, 114–132.

Wijnen, F. (1990). *On the development of language production mechanisms*. PhD thesis. Katholieke Universiteit Nijmegen.

Zamuner, T. (2006). Sensitivity to word final phonotactics in 9- to 16-month-old infants. *Infancy, 10*, 1, 77-95.

Zamuner, T. (2009). Phonotactic probabilities at the onset of language development: speech production and word position. *Journal of Speech, Language and Hearing Research, 52*, 49-60.

Szekely, A., D'Amico, S., Devescovi, A., Federmeier, K., Herron, D., Iuer, G, Jacobsen, T. & Bates, E. (2003). Timed picture naming: Extended norms and validation against previous studies. *Behavior Research Methods, Instruments, & Computers, 35*, pp. 621-633.

Zink, I. & Lejaegere, M. (2003). *N-CDI's Lijsten voor Communicatieve ontwikkeling: korte vormen. Aanpassing en hernormering van de Mac Arthur Short Form Vocabulary Communicative Developmental Inventories*. Acco: Leuven/ Leusden.

English summary

This thesis is about child speech, about the system behind child speech production and about the development and maturation of this system, also referred to as the child speech production mechanism. I investigated its development by looking at the production and at the perception of onset clusters in Dutch toddlers.

The leading question in this thesis is, which part of the speech production mechanism remains "under construction", when a child reduces an onset cluster word (which results in an error with respect to the adult target form). Stated differently, which part of the speech production mechanism is still incomplete, when a toddler produces 'tein', while pointing to a train ('trein'). To obtain an answer to this question I carried out a number of different studies, which are described in chapters 2 to 5 in this thesis.

The study in **Chapter 2** concerns mainly the question whether reduced cluster words are fully reduced or whether in some way they exhibit acoustic traces of the omitted second consonant. For this purpose I analyzed the acoustic characteristics of (minimal) pairs of toddler utterances that differed in the presence or absence of an onset cluster in their target forms. I looked at minimal pairs of onset cluster words vs. singleton target words starting with /Cr/-/C/ and /kn/-/k/, like *brood* /bʀoːt/ 'bread' – *boot* /boːt/ 'boat' and *knip* /knɪp/ 'cut' – *kip* /kɪp/ 'chicken'. The results in Chapter 2 showed evidence for acoustic traces of the omitted consonants.

Even though Dutch children's target /r/ sound in the complex cluster /Cr/ were not perceived by adults, the children tended to realize a rising F2 in the vowel onset, which might be reminiscent of the rising F3 that we see in adult speech. As for target /kn/ words in which the nasal sound was neither perceived by adults, I found that the subsequent vowel did show a moving formant pattern, and perhaps a lower center of gravity. In sum, the data of two-year-old Dutch children showed small acoustic traces of some typical auditory cues of the target consonant in the subsequent vowel. The fact that there is an acoustic trace in the child reduced cluster might suggest that the lexical representation of the cluster word a /CCV/ rather than a /CV/ form.

In **Chapter 3** I have presented a detailed analysis of the acquisition of /Cr/ clusters, by means of descriptive and qualitative analyses of longitudinal data of five children. Seven developmental stages were discovered, but not all children pass through all seven stages, and stages can be overlapping. All children start with the complete and traceless omission of target C_2 (stage 1), like in this utterance of the word "crocodile" by Cato in het fifth session [kɔkɪdju], at the age of two years. The locus of this omission was suggested to be at the level of phonological encoding, where the syllable spellout only recognizes singleton onsets. The next stages consist of an acoustic trace (stage 2) or a substituent vowel following C_1.(stage 3), an example of which would be Cato's [kuokodɪu] in her 8th session, at the age of two years and two months. This stage usually is followed by epenthesis, like in Tirza's utterance of the word 'stairs' *trap* as [tərɑp] in her 16th session at age two years and three months. Here the epenthesis is interpreted as the acoustic effect of immature coordination of the two consonant articulations. A consonant substitution of /r/ in the C_2 position of the onset (stage 5) followed by a realization of epenthesis plus substitution (stage 6) and finally the realization of a mature or not yet entirely mature /r/ (stage 7) are the next stages in the development. An example of stage 7 would be Cato's crocodile uttered as [kʀoːkodɪl] in her 11th session, at the age of two years and four months.

In **Chapter 4** too I looked at the development of cluster words realization. However where in Chapter 3 I used spontaneous speech data, here elicited speech data was used. The experiment involved picture naming (PN), word repetition (WR) and non-word repetition (NWR), performed by four children, where the target forms were real words or nonwords with onset cluster.

The analyses revealed a developmental shift (for three children), from an initial state to a final state. In the initial state performance on NWR outranks performance on both WR and PN, while in the final state performance on all tasks is similar, it has improved and this is especially visible in the PN and WR.

In children's cluster productions, it was observed that the /sC/ and /C/ + liquid clusters show a different timing in development, for two children correct /C/ + liquid clusters appeared first, for one child correct /sC/ cluster appeared first. For one of the four children, Lars the results were unclear. This followed the literature, suggesting an extra-syllabic position for /sC/ clusters and an onset syllable position for the /C/ + liquid cluster. The different syllabic positions clarify the different timing in development.

Lars exhibited performance, which is comparable and poor on all tasks and contrary to the other children this remained the case until the final recordings.

The initial poor performance on the WR and the PN tasks (with respect to the NWR) can be interpreted either as a problem with lexical access (lexical representation) or with phonological encoding. Perception data is needed to tear the two apart.
As soon as the problem experienced in the initial state is solved (either by a better lexical representation or a better phonological encoding), the lexical route could boost the production performance. With respect to the different performance patterns of C + liquid and /sC/ clusters, I cannot draw any clear conclusions about the locus of the problem. Finally, given the persistently poor performance of Lars, I concluded that his speech production mechanism (at least with respect to consonant clusters) was not functioning well. It is very possible that multiple layers of his production mechanism did not function well.

In **Chapter 5** I questioned how detailed the representation of onset clusters is in the child mental lexicon. Do children exhibit different looking behavior when they perceive correctly produced onset clusters as opposed to reduced onset clusters? If so, do they exhibit different looking behavior where C1 is the omitted element, or where C2 is the omitted element? Therefore I examined the perception of correct vs. reduced /sC/ clusters like in the word *stoel* /stul/ 'chair' and /C/ + liquid clusters like in the words *trein* /tʀɛin/ 'train' and *bloem* /blum/ 'flower' clusters by two-year-olds. The results pointed out that two-year-olds exhibit awareness of /sC/ cluster reduction but not of /C/ + liquid cluster reduction. In other words, children looked longer at the picture of a chair (*stoel*) when they perceived the reduced form *toel* as compared to the correct form *stoel*, while they looked equally long at the picture of a train while perceiving correctly and incorrectly produced *trein* and *tein*.

A possible interpretation of the obtained results suggest that the longer looking times at *toel* could evidence for a more stable mental representation of correct /sC/ cluster words as opposed to /C/+ liquid words, in line with the literature on perception of mispronounced words (Swingley & Aslin, 2000, White & Morgan, 2008).

However, my research points to a more probable interpretation of the results, which states that the longer looking times point to a learning effect. The results suggested that in children who have not acquired /sC/ clusters yet this learning effect appears to be stronger than in children who have acquired /sC/ clusters. Children were aware of the difference between 'stoel' and 'toel', while they uttered 'toel' themselves. This might be a sign that they were learning the correct form. In this way the results are asymmetrical. The asymmetry consists in the fact that on the one hand the children who (on the basis of their production) seemed to have stored /CCV/ forms in the mental lexicon, did not show awareness of cluster reduction in perception, while on the other hand, the children who seemed to have stored /CV/ forms (also on the basis of their production), did exhibit an awareness of cluster reduction in perception.

In **Chapter 6** a discussion of the results obtained in Chapters 2 to 5 is offered. First I discussed the nature of the child speech production mechanism in its initial state. However in this thesis the age of interest is when onset clusters are being acquired and I assume that at this age, the speech production mechanism has got the characteristics of the adult mechanism, being formed by a lexical level, phonological level, phonetic level and the articulation level as a final result. In this thesis I offer an account for the maturation of the speech production mechanism according to which it matures in a top-down fashion. As it was said, the levels are already available so what matures, is the representation, which is formed at each of these levels. If we take the development of the production of /Cr/ clusters for instance, we see how first limitations come from the syllable spell out (which is part of the phonological encoding), later they come from the phonological encoding, followed by limitations of the phonetic encoding, when its coordination fails until finally the forms at all levels have matured and an adult-like 'trein' can be uttered.

Samenvatting

Dit proefschrift gaat over kinderspraak, over het systeem dat achter kinderspraak zit en over de ontwikkeling en het rijpingsproces van dit systeem. Dit systeem wordt ook wel het spraakproductiemechanisme genoemd. Ik onderzocht de ontwikkeling van het spraakproductiemechanisme door te kijken naar de productie en de perceptie van woorden die beginnen met twee medeklinkers[1] van Nederlandse peuters.

De hoofdvraag in dit proefschrift is: welk gedeelte van het spraakproductiemechanisme blijft in gebreke, wanneer een kind het begin van een CC-woord reduceert to één C (afwijkend ten opzichte van de volwassen vorm), Met andere woorden, wat is er nog onvolledig aan het spraakproductiemechanisme als een peuter 'tein' zegt terwijl zij *trein* bedoelt. Om deze vraag te kunnen beantwoorden, voerde ik een aantal onderzoeken uit. Deze zijn beschreven in de hoofdstukken 2 t/m 5.

Het onderzoek in **Hoofdstuk 2** zoekt een antwoord op de vraag of gereduceerde CC-woorden (*trein* uitgesproken als 'tein') volledig gereduceerd zijn of dat ze akoestische sporen vertonen van de weggelaten medeklinker. Hiervoor analyseerde ik de akoestische eigenschappen van door peuters uitgesproken woordparen die (minimaal) van elkaar verschillen in de aan- of afwezigheid van twee medeklinkers aan het begin van een woord. Zo onderzocht ik minimale paren van CC-initiële woorden vs. woorden die beginnen met één medeklinker, zoals *brood* /bʀoːt/ – *boot* /boːt/ en *knip* /knɪp/ – *kip* /kɪp/. Uit dit onderzoek bleek dat de gereduceerde CC-woorden niet volledig gereduceerd zijn, maar akoestische sporen vertonen van de

[1] Als ik het heb over 'woorden die beginnen met twee medeklinkers' zal ik voortaan de afkorting CC-woorden gebruiken, van het Engelse *consonant* (medeklinker).

tweede, weggelaten medeklinker. Voor de /r/ uit medeklinker + r clusters bestond het akoestische spoor uit een een stijgende F2[2] in het begin van de klinker, na de eerste medeklinker. Mogelijk is hier sprake van nabootsing van de stijgende F3 die volwassenen in de klinker na een /r/ realiseren. Voor de /n/ uit /kn/-clusters vond ik een bewegend patroon in de klinker na de medeklinker en een lager zwaartecentrum[3].

Alhoewel het spoor dus aanwezig was, bleken volwassenen in een perceptietest geen gebruik te kunnen maken van dit spoor: ze hoorden geen /r/ in de uitingen van peuters, zoals 'tein' waar de /r/ weggelaten was, met achterlating van een spoor, en geen /n/ in gereduceerde /kn/ clusters met spoor, zoals in 'kip' voor *knip*. Kortom, de data van tweejarige Nederlandse kinderen laten in de volgende klinker kleine akoestische sporen zien, typische auditieve aanwijzingen van de volwassen medeklinker. Helaas kunnen volwassen luisteraars geen gebruik maken van deze informatie. Het feit dat er akoestische sporen aanwezig zijn in het door de peuter gereduceerde CC-woord suggereert dat de lexicale representatie van het CC-woord een volledige /CCV/ representatie is en niet een /CV/ representatie is (waar V voor *vowel* (klinker) staat).

Hoofdstuk 3 onderzoekt de verwerving van /Cr/ woorden aan de hand van een beschrijvende en kwalitatieve analyse van longitudinale data van vijf kinderen. Ik ontdekte zeven ontwikkelingsstadia. Het bleek dat niet alle kinderen alle zeven stadia doorlopen en dat sommige stadia kunnen overlappen. Alle kinderen beginnen met uitingen waar de tweede medeklinker, de /r/, weggelaten is en waar geen akoestisch spoor aanwezig is in de klinker

[2] F2 is de tweede formant, samen met de eerste en de derde formant een belangrijke akoestische eigenschap van klinkers.

[3] Zwaartecentrum is *center of gravity*, een andere belangrijke akoestische eigenschap van klinkers.

(stadium 1), zoals in de uiting van het woord *krokodil* door Cato wanneer zij twee jaar is, uitgesproken als [kɔkɪdju]. De locatie in het spraakproductiemechanisme die ik aanwees als bron voor deze weglating is de laag waar fonologische encodering plaatsvindt. Daar vindt syllabificatie plaats, en in dit stadium kunnen alleen lettergrepen gevormd worden met één medeklinker voor de klinker. Er is dus geen plaats voor de /r/, en die verdwijnt. De volgende twee stadia zijn uitingen met een akoestisch spoor (stadium 2) of uitingen waar de /r/ door een klinker wordt vervangen (stadium 3). Een voorbeeld van stadium 3 is woord *krokodil* uitgesproken als [kuokodɪu] door Cato wanneer zij twee jaar en twee maanden is. Het volgende stadium is meestal epenthese, zoals in de uitspraak van het woord *trap* door Tirza wanneer zij twee jaar en twee maanden is, als [tərɑp]. De epenthese wordt hier geïnterpreteerd als een akoestisch effect van de onrijpe coördinatie van articulatoren bij de articulatie van de twee medeklinkers. De volgende drie stadia in de ontwikkeling zijn: een substitutie van de /r/ (stadium 5) of een epenthese plus substitutie (stadium 6) en uiteindelijk de uitspraak van een volwassen of bijna volwassen /r/ (stadium 7). Een voorbeeld van stadium 7 is Cato's uitspraak [kʀoːkodɪl] voor *krokodil*, op de leeftijd van twee jaar en vier maanden.

In **Hoofdstuk 4** gaan we verder in op de ontwikkeling van CC-woorden. Waar er in Hoofdstuk 3 gebruik werd gemaakt van spontane spraak, richt de studie zich hier op uitgelokte spraak. Het experiment was longitudinaal en bestond uit drie taken, plaatjesbenoeming (PN), woordherhaling (WR) en nonsenswoordherhaling (NWR)[4], uitgevoerd door vier kinderen. De kinderen voerden de taken één keer per maand uit, tot het moment dat CC-woorden correct uitgesproken konden worden - of wanneer dat niet gebeurde binnen de tijd die voor het experiment stond, totdat de opnames beëindigd werden. De

[4] PN staat voor picture naming, WR voor word repetition en NWR voor nonword repetition.

analyse toont een ontwikkeling in de uitvoering van de verschillende taken, die ons iets kan vertellen over de ontwikkeling van het spraakproductiemechanisme. In het beginstadium is de uitvoering van de NWR-taak beter dan die van de WR en de PN taken. In het eindstadium is de uitvoering van alle taken gelijk en verbeterd, alhoewel deze verbetering voornamelijk zichtbaar was in de PN en de WR taken.

Uit de uitingen van CC-woorden blijkt dat woorden die beginnen met een /sC/ of een /C/ + vloeiklank (/r/ of /l/), een verschillende ontwikkeling vertonen. De cluster in de /C/ + vloeiklank woorden werd door twee kinderen eerder correct geproduceerd dan die in de /sC/ woorden, terwijl voor één kind die volgorde juist andersom was. Voor één kind was het onduidelijk, omdat hij ook aan het eind van de opnames nog geen één cluster correct kon produceren. Deze resultaten zijn in overeenstemming met de literatuur. Ze suggereren namelijk dat de /s/ van de /sC/ woorden een plaats buiten de lettergreep aanneemt, terwijl beide medeklinkers in de /C/ + vloeiklank woorden een plaats binnen het begin van de lettergreep aannemen, wat het verschil in timing van de ontwikkeling van deze twee type CC-woorden verklaart. Het ene kind begint met het ene type, het andere kind met het andere type.

De matige uitvoering van de WR en PN taken in vergelijking met die van de NWR taak in het beginstadium kan betekenen dat er een probleem met de lexicale toegang tot het woord (de lexicale representatie) is, of dat er een probleem is met de fonologische encodering. Perceptiedata zouden hier uitsluitsel over kunnen geven, maar die ontbreken momenteel. Bij de NWR taak kunnen deze problematische lagen in het spraakproductiemechanisme overgeslagen worden, waardoor de uitvoering op deze taak in het begin beter gaat. De verbetering van de PN en de WR in het eindstadium kan als volgt geïnterpreteerd worden: zodra het probleem dat ervaren wordt in het begin (ofwel met lexicale representatie ofwel met fonologische encodering) opgelost is, zorgt de lexicale route juist voor een betere en meer stabiele uitvoering van de taak. Dit komt doordat er een volledige fonologische representatie gebruikt wordt bij het produceren van woorden. De lexicale route (ofwel de route waarbij de representatie van een woord in het mentale lexicon geactiveerd wordt) is niet aanwezig in de NWR taak omdat onbestaande woorden niet opgeslagen zijn in het mentale lexicon, vandaar dat de uitvoering op deze taak in het eindstadium meer variabel is, en achterblijft bij de andere taken.

Hoofdstuk 5 beantwoordt de vraag hoe gedetailleerd de representatie van de CC-woorden is in het mentale lexicon van het kind. Vertonen kinderen verschillend kijkgedrag als ze correct uitgesproken consonantclusters in woorden horen (trein, stoel) ten opzichte van woorden die worden uitgesproken met een cluster die gereduceerd is op de manier waarop kinderen zelf ook hun clusters reduceren (tein, tul)? En zo ja, is dat zowel het geval bij woorden waarbij de eerste medeklinker van de begincluster wordt weggelaten (dus bij /sC/-woorden zoals stoel-toel) als bij woorden waarbij de tweede medeklinker wordt weggelaten (trein-tein)? Om dit te achterhalen, heb ik het kijkgedrag van tweejarigen onderzocht tijdens het auditief aanbieden van correcte vs. gereduceerde woorden waarbij tegelijkertijd op een scherm twee plaatjes werden aangeboden, één van het doelwoord en een afleider. Het gaat daarbij zowel om /sC/ woorden (zoals ‘stoel’) als om /C/ + vloeiklank woorden (zoals ‘trein’ en ‘bloem’). De resultaten laten zien dat tweejarigen aandacht hebben voor het verschil tussen correcte en gesimplificeerde /sC/ woorden, maar niet voor het verschil tussen correcte en gesimplificeerde /C/ + vloeiklank woorden. Kinderen keken bijvoorbeeld langer naar het plaatje van een stoel als ze de correcte vorm ‘stoel’ hoorden, dan wanneer ze de gesimplificeerde uitspraak ‘toel’ hoorden. Daarentegen keken ze even lang naar de gesimplificeerde ‘tein’ als naar de correcte ‘trein’. Opvallend was dat kinderen even lang keken bij correct 'trein', gereduceerd 'tein' en gereduceerd 'tul', maar veel langer bij correct 'stoel'.

Een mogelijke interpretatie van deze resultaten is dat de langere kijktijd bij correct 'stoel' voortkomt uit een stabielere mentale representatie van /sC/ woorden ten opzichte van /C/ + vloeiklank woorden, die de lexicale toegang ten goede komt. Dit is de verklaring voor langere kijktijden bij correcte woordproducties in de literatuur (Swingley & Aslin, 2000, White & Morgan, 2008).

Door de kijktijdresultaten uit mijn onderzoek te koppelen aan gegevens over de manier waarop de jonge deelnemers aan het experiment zelf de clusters uitspraken, lijkt een andere interpretatie echter meer waarschijnlijk. De langere kijktijd bij correcte productie van /sC/ clusters is dan het effect van een bewustwordingsproces. Uit mijn onderzoek is gebleken dat het effect sterker is bij kinderen die de /sC/ woorden zelf nog gereduceerd uitspraken dan bij kinderen die de /sC/ woorden wel correct uitspraken. Kinderen die zelf 'toel' zeiden, hadden veel aandacht voor de geproduceerde vorm die ook de /s/ bevatte. Mogelijk is dit een signaal dat ze bezig zijn de correcte vorm te leren. De meeste kinderen die aan het experiment deelnamen produceerden zelf als wel /C + vloeiklank/ clusters. De resultaten voor de verschillende clustertypen zijn dus asymmetrisch, en de asymmetrie is waarschijnlijk terug te voeren op hun representaties van clusters in het mentale lexicon. Kinderen vertonen geen aandacht voor het verschil tussen correcte en gereduceerde versies van CC clusters bij woorden die, blijkens hun uitspraak, als /CCV/ zijn opgeslagen (in het experiment gold dit voor woorden met een C + vloeiklank cluster), maar wel bij woorden die, blijkens hun uitspraak, als /CV/ zijn opgeslagen (in het experiment hier gold dat met name voor woorden met een /sC/ cluster). De relatie tussen langere kijktijd en bewustwording moet verder worden uitgezocht.

Hoofdstuk 6 biedt een discussie van de resultaten die verkregen zijn in de Hoofdstukken 2 t/m 5. Als eerste, bespreek ik het karakter van het spraakproductiemechanisme van kinderen in z'n begin stadium. De leeftijd dat in dit proefschrift van interesse is, betreft de tijd wanneer kinderen clusters aan het begin van het woord verwerven. In het proefschrift maak ik de aanname dat op deze leeftijd het spraakproductiemechanisme van kinderen de eigenschappen heeft van dat van volwassenen. Het heeft een lexicaal niveau, een fonologisch niveau, een fonetisch niveau en als resultaat het niveau van de articulatie. Ik bied een verklaring die suggereert dat het spraakproductie mechanisme rijpt op een top-down manier, ofwel van boven naar beneden. Zoals al eerder gezegd, de niveaus van het mechanisme zijn al aanwezig dus wat nog kan rijpen, zijn de representaties die gevormd worden op elk afzonderlijk niveau. We kunnen bijvoorbeeld gaan kijken naar de ontwikkeling van de productie van /Cr/ clusters. Dan kunnen we zien hoe de eerste beperkingen uit de vertaling van de lettergreep structuur komen (wat onderdeel is van de fonologische encoderig), hoe ze vervolgens uit de beperkingen van de fonologische encodering komen, dan uit de beperkingen bij de fonetische encodering (wanneer die in de coördinatie tekortschiet), totdat uiteindelijk de vormen op alle niveaus rijp zijn en het woordje 'trein' op een volwassen manier kan worden uitgesproken.

Curriculum Vitae

Margarita Gulian was born in Sofia, Bulgaria on the 8th of July of 1981. She attended language high school Saint Kiril and Methodius from 1994 to 2000 and studied Spanish philology at Sofia University Saint Kliment Ohridski from 2000 to 2001. She obtained her BA Spanish language and culture in 2004 at the University of Amsterdam. In 2007 she obtained her MA in Linguistics at the University of Amsterdam. The MA was part of a two-year research master, with special focus on second language perception. From 2007 to 2013 she worked on her PhD project "The development of the speech production mechanism in young children; Evidence from the acquisition of onset clusters in Dutch" at Leiden University. The PhD project was part of Claartje Levelt's Vidi project "A psycholinguistic model for phonological development". In 2015 Margarita obtained a teachers degree at the University of Amsterdam and is currently working as a Spanish teacher at Joke Smit College in Amsterdam.