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

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*).

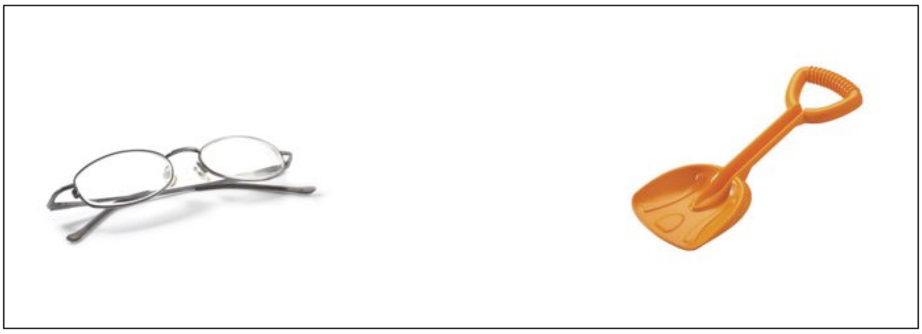

Figure 2: An example from the test phase, where two objects are visible on the screen. In this case the *bril* - *schep* ('glasses' – 'shovel') trial.

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/.

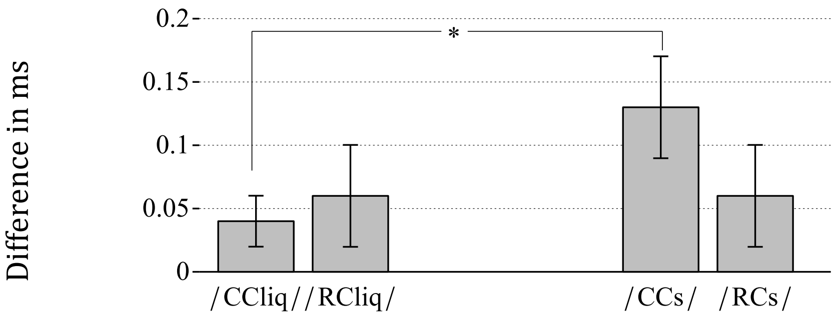

Figure 3: Mean values and SD of the difference PTL score for /C+liq/ words in the correct cluster (CC) vs. /C+liq/ words in the reduced cluster (RC) condition and for /sC/ words in the CC condition vs. /sC/ words in RC condition.

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.

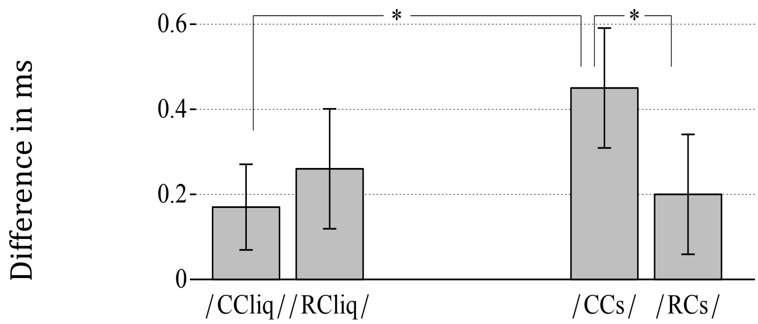

Figure 4: Mean values and SD of the difference LLK score for /C+liq/ words in the correct cluster (CC) vs. /C+liq/ words in the reduced cluster (RC) condition and for /sC/ words in the CC condition vs. /sC/ words in RC condition.

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