Prompted self-repairs in two-year olds

Levelt, C.C.; Brink, E. van den; Karlsson, J.; Weijer, J. van de

Claartje Levelt, Eline van den Brink, and Josefine Karlsson

Prompted Self-Repairs in Two-Year-Old Children

Abstract: Previous work has shown that young children can be prompted to revise an utterance by an unspecific prompt like “hm?” or “what?”. In a game setting this technique is re-introduced with two-year-old children who name pictures on cards that they collect to put on a game-board. The pictures elicit words that start with an onset cluster, like /tr/, /kl/, /st/, which are often reduced to singleton consonants by two-year-old speakers. The prompt is assumed to trigger their self-monitoring system, which could highlight a mismatch between the uttered word – with a reduced cluster – and the intended target – with a full cluster – if this full cluster is indeed part of the young speaker’s segmental representation of the word. We compared the initial production and the prompted repeated productions of 20 children and studied the changes that were triggered by the prompt. Results show that (1) the technique works very well with these young speakers and triggers repeats in most cases, (2) changes involve improvements in most cases, ranging from segmentally more accurate productions and signs of knowledge of an initially absent additional consonant, to full consonant cluster productions, (3) the technique can help to disentangle word-form encoding errors from errors resulting from the stored segmental representation.

Keywords: word production, language development, self-monitoring, self-repairs, toddlers, onset clusters

1 Introduction

In order to produce words, a young language learner needs to simultaneously acquire a lexical representation, a (phonological) grammar, and the means to construct a motor program in real time. It is well-known that it takes time to accomplish this and that during this time children’s word productions often

Claartje Levelt, Leiden University, Leiden University Centre for Linguistics. Leiden Institute for Brain and Cognition
Eline van den Brink, Hanze University of Applied Sciences, School of Business Management
Josefine Karlsson, Örebrö University, School of Humanities, Education and Social Sciences

https://doi.org/10.1515/9783110730098-012

deviate from the adult target. Because of this simultaneous development at different levels, it is not always so clear *what* exactly causes deviations from the target or *where* in the speech production process a deviation arises. Take for example young children's attempts at producing words that start with a consonant cluster, the focus of this paper. In (1) are examples of such productions, of the Dutch child Robin at the age of 1;9, from the CLPF database (Fikkert 1994; Levelt 1994) available in Phonbank (Rose and MacWhinney 2014).

(1) Attempts at word-onset consonant cluster productions

		Target	*Robin (1;9)*
a.	trein *train*	/ˈtrɛin/	[ˈdœn]
b.	stoel *chair*	/ˈstul/	[ˈtuːl]
c.	school *school*	/ˈsχol/	[ˈzoː]
d.	groot *big*	/ˈχʀoˑt/	[ˈzːoˑt]
e.	brommer *moped*	/ˈbʀɔməɹ/	[ˈbuːmɪ]
f.	slapen *sleep*	/ˈslapə/	[ˈpaːpɪ]
g.	clown *clown*	/ˈklɑun/	[ˈtœ̞un]

In all of the productions in (1), in addition to segmental deviations from the target word, the onset cluster of the attempted word is reduced to a singleton consonant. Several accounts have been proposed for this cluster reduction phenomenon, usually in terms of a developing syllable structure grammar, like the presence of a fixed CV(C) syllable template (Menn 1978; Fikkert 1994), or a constraint banning clusters, like*CC or *Complex-Onset (Jongstra 2003; Levelt, Schiller, and Levelt 2000). Usually, the most sonorant segment from the target cluster is omitted to meet the restriction on onset clusters. However, even though a developing syllable structure grammar is a plausible account for cluster reduction in child language, other feasible error sources, like the lexical representation, phonetic encoding, and articulation, should be considered too; the child language data that we study are *utterances* after all.

In this paper we present a study on child language productions, specifically productions of target adult words containing consonant clusters in word onsets, which takes this broader speech production perspective. The study centers around the (re)introduction of an experimental method, self-repair prompting, that will be shown to provide a window on the developing speech production mechanism. Before turning to this method and the resulting data, however, let us shortly introduce the language production model assumed in this work, and the notions of self-monitoring and self-repair that underlie the method.

2 Language Production Model

Our understanding of language production in adults is at an advanced state, and there is general consensus on the big picture of the language production mechanism, based on extensive experimental evidence (Levelt, Roelofs, and Meyer 1999; Indefrey and Levelt 2004; Indefrey 2011; McQueen and Meyer 2019). The model of word production in Figure 1 below, based on the full model of speech production presented in Levelt, Roelofs, and Meyer (1999), focuses on word-form encoding and will be used in the present study.

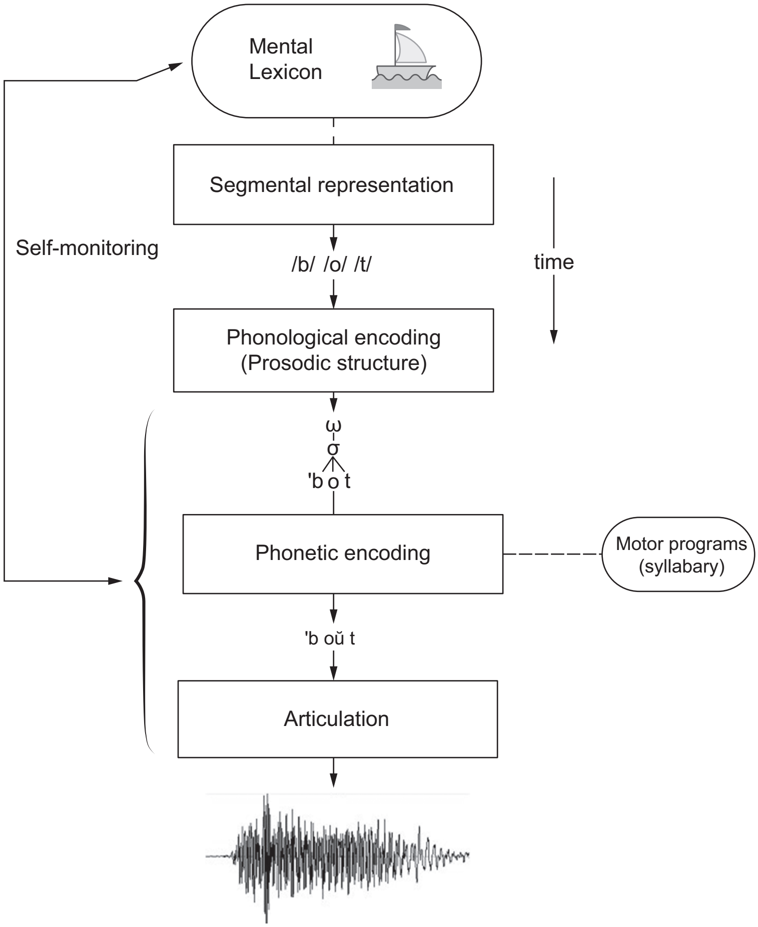

Figure 1: Word-form encoding model (based on Levelt, Roelofs, and Meyer 1999).

Word forms are stored in the *mental lexicon*. In order to prepare the articulatory program that will in the end be executed by the articulators, the phonological representation of a word is retrieved from the mental lexicon. That is, the segmental structure is accessed, as well as word properties that will not be discussed here like the morphological structure and non-predictable metrical structure. In the *phonological encoding module*, the retrieved segmental information will then be spelled out; the segments become available one by one and are grouped into syllables, following both universal and language-specific rules of syllabification. Gestural scores for the syllables are subsequently computed in the *phonetic encoding module*. For mature speakers it has been proposed that gestural scores for (frequent) syllables are stored in a 'syllabary' where they can be accessed directly, so they do not need to be constructed on the fly every time a syllable needs to be articulated. Phonetic fine-tuning is required to smoothen syllable transitions and to adjust the gestural scores to specific contexts. Finally, the gestural score is executed by the *articulatory system*. During the encoding process, speakers monitor their so-called internal speech, that is, the phonological and phonetic representations that are formed. This enables them to correct erroneous representations even before they are executed by the articulatory system.

Up until now it is far from clear how the word production system described here develops in young children. It is reasonable to assume that this system requires a certain period of maturation, just like other cognitive modules (e.g. Dehaene-Lambertz and Spelke 2015), but is it available in this form to the young language learner from the start of word production? And, if fully available, does it perform in the same way? Up until now, hardly any work that takes this perspective on child language data is available (Wijnen 1990, 1992; Levelt 1998; Carter and Gerken 2004; Gulian 2017).

In a fully available word-form production mechanism, the different modules could all form potential *loci* for the deviations we find in word productions of young children. Since we will be focusing on the production of onset clusters, let us take the deviating production [ˈzoː] for *school* /ˈsχol/ (*school*) from (1c) above to illustrate this:

- *The lexical representation*: the word form is *stored* as /zo/, due to perceptual difficulties or problems with the storing of segments. If a target onset cluster is represented with a singleton consonant, only a single consonant will be encoded in the production process.
- *Phonological encoding*: even if the segmental information that is stored in the mental lexicon contains all segments, /s/ /x//o/ /l/, syllabification at the level of phonological encoding could be restricted to CV syllable structures.

This would leave no room for the segments /x/ and /l/, which become 'stray erased' as a consequence. This results in a phonologically encoded syllable /so/, which is subsequently phonetically encoded and receives a gestural score. This scenario subsumes previous phonological accounts of cluster reduction in child language, in which fixed syllabic templates or constraints on syllable structure play an important role.

– *Phonetic encoding*: if phonological encoding has been able to supply the complete syllable /sxo(l)/ – focusing on the onset cluster – this syllable now needs a gestural score. However, constructing a plan with a combination of two consecutive consonantal gestures could be problematic: the gestural scores for [s] and [o] could, for example, overlap, largely masking the score for [x]. The listener is therefore not aware of a [x] sound in the production, but an acoustic analysis could reveal a trace of [x] in the vowel.
– *Articulation*: this scenario is very much like the one for *phonetic encoding*, except that the source of the cluster reduction now lies in the execution of the gestural score rather than in the gestural score itself. The execution of [x] could be masked by an articulatory overshoot of [s] (Shaw et al. 2011).

What kind of information could be used to determine the source of specific deviations and to find out about the developmental state of the speech production mechanism? We could look at the nature of deviations from adult target words in longitudinal child language data. In case some deviation is stable for a period of time, that is, the same error shows up every time a child attempts to produce a specific word, it is likely that the source for the error lies either in the lexical representation or in the phonological encoding module – constrained by the developing phonological grammar. More variable errors are expected if the source lies in the phonetic encoding module or in articulation. We could get more information about the level of segmental detail in young speakers' lexical representations by performing perception experiments, e.g. preferential looking (Swingley 2003, 2005; White, Morgan, and Wier 2005). Acoustic analyses of the produced words could reveal if target segments that seem to be missing from the produced form are indeed completely absent or turn out to be partly realized, revealed by acoustic traces in the signal that are hard to pick up by ear (Gulian 2017). Finally, ultrasound measures could help determine the motor execution abilities of the child (Gibbon 1999). In the present study, however, we will use information provided by a production experiment that revives the method of inducing self-repairs, as used in work by, among others, Gallager (1977) and Käserman and Foppa (1981).

3 Self-Monitoring and Self-Repairs

Self-monitoring entails comparing the encoded, or articulated speech form – here we focus on words – to the intended form, i.e. the perceptual standard of a speaker. If the speaker detects a mismatch between these forms, a correction is made. Self-repairs, or the absence thereof, can provide an important source of information on the state of the production mechanism. First of all, it tells us if children pay attention to their own speech, that is, if the self-monitoring mechanism is in place (see Figure 1) and functioning. Slobin (1978), based on an observation of his daughter at the age of 3;1,[1] remarks that self-monitoring is probably relatively late to develop. However, Jaeger (2005) reports on spontaneous phonological self-repairs in two-year olds acquiring American English, as in: "A blue [pʰa.wi] . . . no . . . [wa.wi.pʰap] (lollipop)" by Alice, 2;4.2. Of the 198 errors recorded by Jaeger in this age group, 38.5% were self-corrected (Jaeger 2005). Gallagher (1977) is to our best knowledge the first study using prompts to induce revisions. The study involved 18 subjects between 20 and 29 months old. In a one-hour play and read session with individual children, an unspecific prompt "what?" was uttered 20 times by the experimenter, independent of the correctness of the child's utterance. These prompts led to revisions in 77% of the cases. Gallagher observed a high frequency of phonetic changes in the youngest group, like in the following interaction. *Child*: "He kit ball" *experimenter*: "What?", *child*: "He kick ball." These phonetic revisions, although they did not always result in the correct target form, were consistently unidirectional, i.e. in the direction of the adult model. According to Gallagher this showed that the child's reaction to "what?" is not merely an attempt to repeat his or her utterance, with some accidental segmental variability due to his or her immature phonological system. Gallagher concludes that *"[. . .] the phonetic changes are revisions within a language system so primitively organized that other structural options are not as readily available*." The unidirectionality of phonetic revisions by these young children was also found by Käsermann (1980), who showed that sounds in words were changed more often towards an adult model than away from it in situations where the mother showed signs of incomprehension. From both these studies it can be concluded that self-monitoring is functioning in children younger than 3 years old. In Käsermann and Foppa (1981) the assumption is, indeed, that the revision behavior of even these young children is guided by their awareness of standard word-forms, and they designed a study to investigate what kind of standards

1 "If her verbal formulations are not at once understood, she lies on the floor and cries or screams – but doesn't attempt to reformulate her statement" (Slobin 1978: §12).

determined children's systematic revisions. In their study, again a 1 hour play and read session, with 4- to 5-year old children, they intervened now and then with an unspecific "*hm?*," which they call a "modification-provoking non-understanding signal." Their hypothesis was that both the occurrence – does the child reformulate or not – and the type of revision a child would make would tell us something about their postulated linguistic standards of correctness and completeness. For this purpose, the utterance before the "hm?" was compared to the utterance prompted by "hm?". Since the children in this study were older, the focus of this study was on grammatical structures rather than on phonological segments. An important conclusion of the study, however, is that triggered repairs are dependent on the existence of certain linguistic standards, that is, on the child's present linguistic knowledge.

Thus, a child's utterance prompted by "hm?" or "what?" can tell us whether an initial, erroneous utterance is or is not in accord with the perceptual standard of this developing speaker. This is important to know, especially in the case of young language learners, who are simultaneously developing competence at different levels. If the initial, deviating, production is in accord with the speaker's perceptual – phonological – standard, suggested by a prompted repetition that does not contain any segmental changes, the error locus is probably situated in the lexical representation, while if it is not in accord with the perceptual standard, suggested by a revised utterance with some segmental change, the source of error should be sought in the form-encoding mechanism. The ability to make self-repairs indicates that children have representations of language forms that are more developed than their productions reveal, but not necessarily adult-like. Self-monitoring thus appears to be an important requisite for language development: only by comparing one's own production to his or her perceptual standard the learner can become aware of discrepancies that necessitate changes in either the lexical representation or the form-encoding system (Clark and Andersen 1979).

In addition to inducing self-monitoring the "hm?" prompt can also be understood as a form of repetition priming (Wheeldon and Monsell 1992; Jacobs et al. 2015; Tsuboi, Francis, and Jameson 2020). In studies with adult speakers, it has been found that repeated productions are facilitated and can have shorter durations. Repetition priming studies with adults have not made use of a prompt but have combined reading or listening to a prime with a subsequent picture-naming task. Up until now it has not been possible to robustly assess the accuracy of picture-naming after a produced prime, due to very low error rates in the primes produced by adult speakers. Shorter durations of primed repetitions were only found when the prime had been overtly articulated (Jacobs et al. 2015), and this was thought to be the effect of priming of the production component through the processing of auditory feedback. According to Tsuboi, Francis, and Jameson (2020) all processes that are not overlearned are susceptible to the repetition priming effect, and more difficult processes will show stronger repetition priming effects. While in the case of adult speakers the articulation of phonological sequences can, in most cases, be called ‘overlearned’, this is certainly not the case for the developing speakers in our study who are targeting consonant clusters. If the mechanism functions like in adult speakers, we can thus expect a strong repetition priming effect, i.e. possibly more accurate and faster post-prompt productions. However, the effect of more accuracy in the post-prompt repetition is at least partially dependent on the segmental information available in the lexical representation.

4 Experiment

Inspired by the work discussed above, we set out to use the unspecific question method in a more restricted way to study the relation between produced form and perceptual standard in the case of consonant clusters in word onsets. Since previous work had shown that the method worked with children as young as 20 months old, we could target the age group that is still struggling with consonant clusters in production, namely children between 24 and 30 months old. This would provide the opportunity to find pairs of initial incorrect productions and induced self-repairs. Below we describe the experimental method, and we will then present two types of analyses of the data, a more quantitative analysis and a more qualitative analysis of a set of systematic repairs, which tells us more about the locus of the initial error in the speech production mechanism.

5 Method

5.1 Participants

We analyzed the data of 20 children from different age groups. In the younger age group there were four boys and six girls between 23 and 25 months old, with a mean age of 24 months. In the older age group there were four boys and six girls between 27 and 31 months old, with a mean age of 30 months. Two more recordings are not analyzed here: data of one 24-month-old girl was not included because it was difficult to recognize individual segments in her speech, and data of one 30-month-old girl was not included in the analysis because her younger sister participated in the experiment, and it was hard to keep their voices apart in the recording. All children had at least one native Dutch-speaking parent. Parents were asked about their children's hearing-, language- and cognitive development; no parent reported any abnormalities. All children were familiar with the pictures-to-be-named during the recording session as the pictures had been sent home to the parents a week before and parents had been asked to show the pictures to their child and encourage the child to name the pictures.

5.2 Stimuli

The test stimuli were Dutch words starting with consonant clusters (/bl, br, dr, kl, kn, kr, fl, xl, vl, vr, sx, sl, sn, sp, st, tr/). One third of the words that could be elicited consisted of phonologically less complex control words, in order to prevent possible frustration with the task. All the words that were used as stimuli had been produced by toddlers in the Dutch-CLPF corpus (Levelt 1994; Fikkert 1994) available in PhonBank (Rose and MacWhinney 2014), indicating that the chosen stimuli were not beyond the possibilities for children at this age. For a list of all the words that could be elicited see Appendix 1.

5.3 Procedure

The experiment was conducted in a quiet room in the Babylab at Leiden University. In order to motivate the children to produce words, and to get a natural setting for self-repair prompts, the test was designed as an interactive game.[2] The child was encouraged to collect nine cards with pictures in color, which fitted on a 3x3 game-board that contained the same nine pictures in black and white. There were 7 different game-boards to choose from, with pictures of well-known animals, household objects, pieces of clothing, or food items. Both the experimenter and the child had a game-board and a stack of cards and they had to ask each other for the card they wanted to fit on the game-board. The main purpose of the game was to elicit specific words, so for some children the card collecting game had to be slightly modified, e.g. some children wanted to collect the cards in the card box instead of on the game-board. When the child named a picture of a word starting with an onset cluster, the experimenter pretended that she had not understood the child and uttered a non-specific request for repetition such as "hm?". If the child did not react to this prompt the experimenter more specifically

2 Thanks to Frank Wijnen for suggesting this!

asked the child to repeat the word by asking “Can you say it again?” or “What did you say?”. There were seven different sets of nine pictures and the aim was to prompt minimally four repeated utterances per set, which was not always possible. All sessions were audio-recorded with a mobile digital recorder (M-AUDIO microtrack II) using a table microphone (RØDE NTG-2 microphone and NC fxx PROEL cable).

5.4 Analysis

All recordings were entered in Phonbank. For the analysis we focused on word-pairs that were produced before and after the prompt, meeting the following criteria: (1) the target word had to contain an onset cluster, (2) the repair could not be a direct imitation of the experimenter’s utterance – since this would not show the self-monitoring effect – and (3) the audio files of both utterances – the original utterance and the repeat – had to be clear, i.e. free from intervening voices or other noise. The data set that fulfilled the above criteria consisted of 236 word-pairs – out of 314 initial productions that were followed by a repetition – produced by the 20 children described above, ranging from 4 to 27 word-pairs per child. Only in 21 cases a prompt did not trigger a repetition.

The utterances were phonetically transcribed by two of the authors, making use of Praat (Boersma and Weenink 2020) to help focus on details of the productions and to compare the pairs of utterances directly.[3] In addition, the durations of the productions before and after the prompt were measured. This was done to check whether, especially in cases where no segmental or structural change could be detected in the prompted production, the repetition priming effect would be noticeable in a faster repeated production after the prompt.

We compared the phonetic transcriptions of each before-after-prompt word-pair, focusing on the realization of the target consonant cluster, and determined whether or not a segmental change had occurred. Regarding consonant cluster development, detailed research by Gulian (2017) had revealed 7 stages in the production of target /Cr/ onset clusters: (1) Complete cluster reduction to a single consonant, (2) Cluster reduction to C_1, with an acoustic trace of C_2 in the following vowel, (3) Cluster vocalization, i.e. production of C_1 and vocalization of C_2, (4) Epenthesis between C_1 and C_2 substitute, (5) C_1+substituted C_2 without epenthesis, (6) Epenthesis between C_1 and correct C_2, (7) Correct cluster realization. While

3 The data are available through CHILDES/Phon and OSF (Data repository of “Prompted Self-repairs in Two-year-old Children”, https://osf.io/kdqa3/).

these stages were not necessarily all encountered in the developmental data of the individual children studied by Gulian, subsets of the stages would always appear in this order. These stages – except for stage 2, because our data was too sparse to robustly establish the presence of acoustic traces – formed the basis to determine whether the prompted repetition involved an improvement or a deterioration in the realization of the target cluster compared to the initial spontaneous word production. In order to classify all the changes we encountered, it was necessary to add some substages; in two cases the target cluster had been completely deleted in the initial production and was improved to a reduced cluster after the prompt. We therefore added a Stage Ø. The single consonant in the cluster reduction stage, Stage 1, could either be the target C_1, the target C_2, or a substitute consonant (C_{sub}). Some of the post-prompt productions involved a change from one type of single C to another type of single C, where a change from C_{sub} to either C_1 or C_2 was regarded as an improvement. Finally, in three cases the production changed from a single C to a very long version of this C. This was interpreted as a variant of the vocalized cluster, Stage 3 in Gulian (2017). This brings us to the following set of stages in the production of target onset clusters:[4]

Stage 0: Ø	Stage 2b: C::
Stage 1a: C_{subst}	Stage 3: CVC_{subst}
Stage 1b: C_1	Stage 4: CC_{subst}
Stage 1c: C_2	Stage 5: $CVC_{correct}$
Stage 2a: CV	Stage 6: $CC_{correct}$

Finally, changes could also affect other parts of the word. These were not analyzed in detail, but just classified as either improvements or deteriorations.

6 Results and Discussion

6.1 Aggregated Findings

Of the 236 word-pairs, the vast majority, 166 (70.3%) contained some change in the repeated production (see Table 1 below). In 109 of these cases (46%) a change took place in the production of the target cluster, of which more than half, 62 (56.9%), were clear improvements, in the sense that the production of the target cluster in the repetition belonged to a more advanced developmental stage than

4 One category that is still unaccounted for in this system is a single C that merges feature information from both target consonants, like [m] for /br/.

the one produced initially. In 18 cases (7.6%) the repeated production was even corrected from a singleton onset in the initial production to a (correct) cluster. In 6 cases the target cluster remained reduced to a single segment but changed to the other consonant from the target cluster. In 7 cases a single substituted consonant changed to one of the target cluster consonants. In 25 cases (22.9%) the production of the cluster deteriorated after the prompt, that is, the speaker reverted to a production from a less advanced developmental stage after the prompt. In the remaining cases some segmental change occurred, but this could neither be classified as an improvement, nor as a deterioration.

In 92 out of the 236 repeated productions (39%) there was a change in a different part of the word (too), with 40 (43.5%) clear improvements and 21 (22.8%) clear deteriorations. Deteriorations, in this case, involved for instance the dropping of a coda consonant that was correctly produced before the prompt, the production of an additional consonant or vowel – not related to the production of the target cluster and also not present in the target word – or a change from a correct to an incorrect consonant or vowel. In 18 cases the sentence structure (also) changed after the prompt: words were added, an article was added or removed, or the word was changed from or to a diminutive or plural form.

In 67 (28.4%) of the 236 word-pairs there was no segmental or structural change at all in the repeated production – but sometimes the repeated utterance was either produced in a softer voice, or it became louder or more emphatic.

In terms of duration, in 78 (74.3%) out of the 105 cases in which the before and after production had similar (numbers of) segments, and were produced in similar sentence positions, the duration of the repeated production was shorter (mean 211ms shorter, range 3ms–1047ms), while in 25 (23.8%) cases it was longer (mean 191.2ms longer, range 10ms–1412ms).

The 24-month-olds produced more repetitions that contained changes than the 30-month-olds (in 77.4% vs. 62.5% of the repetitions), and these changes also involved the production of the cluster more often (50.8% vs 38.4% of the repetitions). Almost 70% of the changes in the repeated production of the cluster were improvements in the 30-month-olds, while for the 24-month-olds changes were improvements in 54% of the cases. A higher percentage of the changes in the repeated production of the cluster involved deteriorations for the 30-month-olds than for the 24-month-olds (27.9% vs. 17.5%). These numbers can be understood when we look at the initial productions of the target clusters in the two groups. Here a major, and expected, difference was found in the percentage of onset clusters produced (correctly) in the initial production: 13.7% (4.8%) for the 24-month-olds versus 42.9% (29.5%) for the 30-month-olds. For the 24-month-olds there is thus more room for improvement, while the 30-month-olds have a higher risk of deteriorations in their cluster production. The higher percentage of changes in general in the 24-month-olds could signify a less robustly operating word-production mechanism. However, this requires further study.

From the numbers above we can conclude that the method successfully triggers self-monitoring; the prompt triggers a repetition, which in the majority of cases involves a change. In comparable before-after pairs, the repeated word was also produced with a shorter duration in most cases, a repetition priming effect. In 56.9% of the changes in repeated target cluster productions, and in 43.5% of the changes elsewhere in the word the rendered production was more accurate, showing that – at least in these cases – self-monitoring indeed functioned in a mature way, as it signified errors and led to self-repairs. Let us now turn to the data in more detail.

7 Changes in the Prompted Productions of Target Clusters

7.1 Overview

In Table 1 below all the different types of changes in the repeated production of the target onset cluster, i.e. improvements, denoted by x, and deteriorations, denoted by o, are tallied. The grey cells denote "no changes" which we ignore here.

Three main observations can be made. The first one is that most improvements occur when the initial spontaneous production is from Stage 1a or 1b. While at first sight this is only logical, because there are simply many ways in which a Stage 1 production can be improved, it is actually highly informative with respect to the question whether the initial productions were produced according to the perceptual standards of the participating developing speakers. The answer in these cases is "no": these speakers can correct themselves, triggered by the prompt, indicating that more information is stored in their lexical representations than the initial productions would lead us to believe.

The second observation is that pre-prompt productions belonging to Stage 2 do not appear to change very often. At this stage the speaker is clearly targeting more than a single consonant, pointing to a lexical representation that contains information about both consonants from the target cluster. The fact that we do not find many changes could of course simply be an effect of the fact that we have a constrained set of data from this stage. However, it could also point to a production solution that developing speakers feel comfortable with for some time: both

Table 1: Status of the productions of target clusters of all 20 participants before and after the prompt. Cells contain production tokens. An x denotes an improvement, an o denotes a deterioration in the repetition.

before → / **after ↓**	**Stage 0:** ∅	**Stage 1:** C_{sub}	**Stage 1:** C_1	**Stage 1:** C_2	**Stage 2:** CV	**Stage 2:** C::	**Stage 3:** CVC_{subst}	**Stage 4:** CC_{subst}	**Stage 5:** $CVC_{correct}$	**Stage 6:** $CC_{correct}$
Stage 0: ∅			o							
Stage 1: C_{sub}			ooo					o		
Stage 1: C_1	xx	xxxxx			oo		oo	oo	o	ooo
Stage 1: C_2		x	xxxxx					oo		
Stage 2: CV		x	xxxxxx						o	
Stage 2: C::			xxx							
Stage 3: CVC_{subst}					xx					
Stage 4: CC_{subst}		xxxx	xxxxxx				xxxxx		o	o
Stage 5: $CVC_{correct}$			x				x	xx		o
Stage 6: $CC_{correct}$			xxxx	xx				xxxxxxx	xxxxx	

consonants are realized, while the unmarked core syllable structure CV is still observed: a win-win situation. This hypothesis awaits further testing.

The final observation is that when the initial production is from Stage 3, 4 or 5, changes go both ways. Here the prompt either leads to a more accurate production, or it triggers developing speakers to revert to productions from an earlier stage with which they have more experience. This could be related to speakers' familiarity with the particular stage. In case speakers have just entered the stage, they might revert to a more comfortable previous stage after the prompt, while if speakers are about to reach the next stage the prompt might function as a boost to attain this next level production. This could not be tested with the present set-up of the study. To do this, experiments with prompted repetitions should be repeated over time with a group of developing speakers, while their spontaneous speech is collected at regular intervals over this longer period as well. In this way it can be determined whether a speaker has just entered a specific developmental stage or has been there for a while, and whether this affects the type of change in a prompted repetition.

Let us now discuss actual examples of word productions before and after the prompt.

7.2 Hidden Knowledge

There are several types of changes that reveal that a young speaker who produced a reduced cluster in the word production before the prompt, has more detailed information about the target cluster available in the lexical representation. That is, there is knowledge, initially hidden, about the full target cluster that is revealed by the post-prompt production. Below in (2) are examples of a C_1-C_2 swap: the speaker produced one of the two consonants, or a regular substitute for that segment, of the target cluster before the prompt and reverted to the other consonant, or a regular substitute for that segment, in the production after the prompt.

(2) Hidden knowledge I: C_1-C_2 swap

	Target	*Gloss.*	*IPA target*	*Before prompt*	*After prompt*	*Participant*
a.	stoel	chair	/stul/	[tˢʊ]	[sʊ]	N24
b.	broodje	bun	/brotjə/	[doʃə]	[loʃə]	H24
c.	kraan	faucet	/kran/	[kan]	[lan]	A24
d.	slang	snake	/slɑŋ/	[tˢɑn]	[lan]	D24

In (3) are examples of cases where the speaker again produced only one of the two consonants, or a regular substitute for that segment, of the target cluster before the prompt, but added (partial) information about the other consonant in the repeated production. In these examples the initial part of the repeated utterance is still CV, but now the V is either a vocalized version of the target C_2, or an epenthesized vowel followed by C_2 – or a consonant substituting this consonant.

(3) Hidden knowledge II: A vocalized or epenthesized cluster after the prompt

	Target	*Gloss*	*IPA target*	*Before prompt*	*After prompt*	*Participant*
a.	broek	trousers	/bruk/	[bux]	[puu:x]	A24
b.	fles	bottle	/flɛs/	[fɛs]	[fɪɛs]	N24
c.	druifje	grape (dim)	/drœyfjə/	[ʋœyçʲə]	[tĭœysʲə]	H24
d.	brood	bread	/brot/	[poˑt]	[puoːt]	T24

In (4), finally, are cases where a speaker produced one of the two consonants, or a regular substitute for that segment, of the target cluster before the prompt and a full-fledged consonant cluster after the prompt.

(4) Hidden knowledge III: full-fledged clusters after the prompt

	Target	*Gloss*	*IPA target*	*Before prompt*	*After prompt*	*Participant*
a.	sleutel	key	/sløtəl/	[toto]	[klotoŭ]	A24
b.	sleutel	key	/sløtəl/	[lɔtɔ]	[tlotɔ]	L30
c.	sloffen	slippers	/slɔfə/	[s̠ˑafɔ]	[s̠ˑl̠æ̆ufɛ]	C24
d.	slang	snake	/ slɑŋ/	[θɑŋ]	[tlɑŋ]	H24
e.	klok	clock	/klɔk/	[lɔk]	[klɔk]	R27
f.	broek	trousers	/bruk/	[puːk]	[bʀuk]	N30
g.	schoen	shoe	/sxun/	[xu]	[sxun]	L30

The target clusters in (4a–e) all have /l/ as the target C_2 and a [Cl] cluster is produced in the repeated productions. The participants who produced these remarkable improvements all had other instances of [Cl] realizations in their recorded data. For example, C24 produced [pɬ::œt] for target *brood* (bread), substituting target C_2 /r/ with [l], but in this recording also produced epenthesized clusters, like in [kɑlʌk] for *klok* (clock), and reduced target /Cl/ clusters, like in [pa:] for *bloem* (flower). Participant C24 is clearly working on the realization of this specific cluster type and this also applies to the other participants in (4a–e). There is, thus, some articulatory experience with the production of a [Cl] cluster, and the repetition seems to provide the speaker with the opportunity to supply the information that was missing from the original reduced pre-prompt production. The same applies to the example in (4f), but here involving /r/ as the target C_2; instead of correct [Cl] clusters, N30 produced several instances of correct [Cr] clusters in initial productions. The example in (4g), finally, involves a target /sx/ cluster, produced by L30 who already produced both target /Cl/ and /Cr/ clusters correctly.

The pairs of productions discussed in this section clearly show that the source of the reduced clusters in the pre-prompt productions does not lie in the stored lexical representation of the words. That is, it is not the case that the cluster reduction in the pre-prompt production results from a lexical representation that contains a single consonant instead of a consonant cluster at the word onset. The post-prompt production demonstrates that information about both consonants of the cluster in the target word is – at least partially – available in the lexical representation. The source of the deviating production must therefore lie at the level of phonological encoding or at the phonetic-articulatory level.

In the case of the before-after productions in (2) and (3) above, constraints at the phonological encoding level could play a role. The production of the target cluster structurally deviates in both the production before and the production after the prompt, in the sense that instead of CCV sequences only CV sequences are realized. If syllables with complex onsets cannot be construed in the phonological encoding module, due to a constraint that only allows for the construction of CV syllables, then a /C_1C_2V/ sequence retrieved from the lexicon must be accommodated in some way: by encoding only one of the consonants, C_1V or C_2V, by vowel epenthesis, $C_1\mathit{V}C_2V$, or by assigning the C_2 to the V position, which results in a vocalized cluster, CVV, as in 3a–d above.

In the case of the before-after productions in (4) above, however, the error locus must lie at the articulatory-phonetic level. The fact that a full cluster is produced after the prompt shows that the syllable spell-out at the phonological encoding level is not constrained to single onset positions. Two scenarios can be considered to account for the production variation illustrated in (4). In one, two motor plans are activated for the target /CCV/ sequences, the old [CV] plan and a new [CCV] plan. Since the speaker has more experience with the old plan, this will be activated above the threshold level first and is executed. The activation of the executed old plan subsequently drops, and the prompt and resulting self-monitoring now provide the new plan – that has still retained some activation – with the opportunity to cross the threshold level for execution. In the second scenario the locus of the variation lies in the execution of a single motor plan, containing the gestural scores for both consonants. In the initial execution of this plan the motor coordination of the gestures is off, resulting in the – perceptual – masking of one of the consonants. The prompted self-monitoring correctly identifies the problem, and in the repetition the coordination is improved. At this point, and on the basis of the sparse data available in this study, it is not possible to decide which of the two scenarios is most likely.

7.3 Full Knowledge-Lack of Knowledge

In 27.5% of the repetitions after the prompt no noticeable segmental change occurred. For some children the repeated production was softer, or even whispered, while for other children it was louder, or more emphatic. While there might be all kinds of reasons why nothing changed, like a lack of attention or increasing fatigue, here we consider two possible causes related to the development of the speech production mechanism.

One child was clearly bored with the repetition procedure.[5] This child was one of the 30-month-olds and produced almost all the target clusters correctly in the initial spontaneous production. In this case, the prompted self-monitoring procedure comes up with nothing to correct or improve. The production mechanism appears to be fully developed: the lexical representation is fully specified, word-form encoding applies accurately – the motor program might have been stored in the syllabary – and there is apparently enough articulatory experience with the motor program to execute it flawlessly every time.

On the flip side, some children produced mostly reduced clusters in the initial production and repeated this without a change after the prompt. In these cases, too, the prompted self-monitoring appears to come up with nothing that can be corrected or improved. However, because both the initial production and the repetition deviate in the same way from the adult target, in these cases knowledge seems to be lacking somewhere in the mechanism. The two most likely sources are the lexical representation and the phonological encoding module. The reduced cluster in both productions either originates in the lexical representation of the target word, which would lack the specification for one of the onset consonants, or results from the presence of constraints on complex onsets at the level of phonological encoding. Alternatively, the self-monitoring system could be the source, leaving errors unnoticed. However, in our data all children produced at least some self-repairs, making this option less likely. In one case, L30, the mechanism did seem to be particularly immature, as 9 out of 11 initial productions contained reduced clusters, 6 prompts led to unchanged repetitions and 4 out of 7 repetitions had longer rather than shorter durations. Again, we would need more information to decide on the exact locus, and this could differ for individual children. For example, an additional perception experiment could, in a future study, determine whether both consonants from target words with onset clusters are specified or not.

5 In one case E30 emphasized this by adding the word "poep" (poop) to the repeated word!

8 Conclusions

Two main conclusions can be drawn from this study. First, the method, using prompts to trigger repetitions, is highly suitable to perform with two-year-olds and provides highly informative data. The prompt almost always leads to a repetition and appears in most cases to successfully trigger self-monitoring, resulting in self-repairs. These repairs provide us with a window on the developing speech production mechanism, and we can conclude that the method adds to our understanding of children's development of language production, by disentangling representational from form-encoding immaturity.

Second, the fact that the source of deviating productions could be located at the lexical representation, at the level of phonological encoding, or at the level of phonetic encoding/articulation, demonstrates that studies of child language development cannot do without a detailed speech production model.[6] Self-monitoring appeared to function, and shorter durations of the repeated productions, an effect of repetition priming, were found. Together with the different sources for error in the model that could be established, this seems to indicate that the basic layout of the word-form encoding part of the speech production mechanism is available to the developing speakers.

Finally, we hope that the promising prompted repetition method re-introduced here will help to revive the use of production experiments to study, in particular, early typical child language development. Especially when performed in tandem with perceptual or articulatory studies they can provide us with still much needed and robust information on the developmental state of the language production system.

Appendix I: Stimuli

Test items (consonant cluster in onset)	***Distractors (no consonant cluster in onset)***
Animals	
schaap *sheep*	poes *cat*
vlinder *butterfly*	aap *monkey*
slang *snake*	kip *chicken*

6 And in order to fully understand the contribution of the lexical representation to production deviations, the same necessity applies to a model of speech perception.

(continued)

Test items (consonant cluster in onset)	*Distractors (no consonant cluster in onset)*
krokodil *crocodile*	beer *bear*
spin *spider*	
Clothes	
broek *trousers*	jas *coat*
knoopjes *buttons*	sokken *socks*
trui *sweater*	pet *cap*
schoen *shoe*	
slab *bib*	
sloffen *slippers*	
In the house	
stoel *chair*	bed *bed*
klok *clock*	telefoon *phone*
kraan *faucet*	tafel *table*
trap *stairs*	kast *cupboard*
	bad *bath tub*
Transportation	
trein *train*	auto *car*
graafmachine *excavator*	fiets *bicycle*
slee *sled*	boot *boat*
vliegtuig *airplane*	
tractor *tractor*	
vrachtauto *truck*	
Food	
snoepjes *candy*	appel *apple*
druiven *grapes*	taart *cake*
brood *bread*	kaas *cheese*
fles *bottle*	banaan *banana*
spaghetti *spaghetti*	
Play	
glijbaan *slide*	ballon *balloon*
trommel *drum*	konijn (knuffel) *rabbit*
schommel *swing*	pop *doll*
knuffel *cuddle toy*	puzzel *puzzle*
blokken *blocks*	boek *book*
trompet *trumpet*	

(continued)

Test items (consonant cluster in onset)	*Distractors (no consonant cluster in onset)*
Other items	
bril *glasses*	kam *comb*
sleutel *key*	kaars *candle*
bloem *flower*	hamer *hammer*
ster *star*	ijsje *ice-cream*

References

Boersma, Paul & David Weenink. 2020. Praat: Doing phonetics by computer [Computer program]. http://www.praat.org/.

Carter, A. & L. Gerken. 2004. Do children's omissions leave traces? *Journal of Child Language* 31. 561–568.

Clark, E. V. & E. S. Andersen. 1979. Spontaneous repairs: Awareness in the process of acquiring language. *Papers and Reports on Child Language Development* 16. 1–12.

Dehaene-Lambertz, G. & E. Spelke. 2015. The infancy of the human brain. *Neuron* 88(1). 93–109. DOI: 10.1016/j.neuron.2015.09.026. PMID: 26447575.

Fikkert, P. 1994. *On the acquisition of prosodic structure*. Leiden: Leiden University Doctoral dissertation. (HIL Dissertation Series 6). The Hague: HAG.

Gallagher, T. M. 1977. Revision behaviors in the speech of normal children developing language. *Journal of Speech, Language, and Hearing Research 20*(2). 303–318.

Gibbon, F. 1999. Undifferentiated lingual gestures in children with articulatory/phonological disorders. *Journal of Speech, Language, and Hearing Research* 42. 382–397.

Gulian, M. E. 2017. *The development of the speech production mechanism in young children: evidence from the acquisition of onset clusters in Dutch*. Leiden: Leiden University LOT Dissertations in Linguistics.

Indefrey, P. 2011. The spatial and temporal signatures of word production components: A critical update. *Frontiers in Psychology* 2. Article 255. DOI: 10.3389/fpsyg.2011.00255.

Indefrey, P. & W. Levelt. 2004. The spatial and temporal signatures of word production components. *Cognition* 92. 101–144.

Jacobs, C., L. Yiu, D. Watson & G. Dell. 2015. Why are repeated words produced with shortened durations? Evidence from inner speech and homophone production. *Journal of Memory and Language* 84. 37–48.

Jaeger, J. 2005. *Kids' Slips. What Young Children's Slips of the Tongue Reveal about Language Development.* Mahwah, NJ/London: Lawrence Erlbaum Associates.

Jongstra, W. 2003. *Variation in reduction strategies of Dutch word-initial consonant clusters*. Toronto: Toronto University Doctoral dissertation. Toronto University Working Papers.

Käsermann, M. 1980. *Spracherwerb und Interaktion*. PhD Dissertation. Bern: Hans Huber.

Käsermann, M. & K. Foppa. 1981. Some determinants of self correction: An interactional study of Swiss-German. In W. Deutsch (ed.), *The Child's Construction of Language*, 97–104. London: Academic Press.
Levelt, C. 1994. On the acquisition of Place. Leiden: Leiden University Doctoral dissertation. (HIL Dissertation Series 8). The Hague: HAG.
Levelt, W. 1998. The genetic perspective in psycholinguistics. Or: where do spoken words come from? *Journal of Psycholinguistic Research* 27(2). 167–180.
Levelt, C., N. Schiller & W. Levelt. 2000. The acquisition of syllable types. *Language Acquisition* 8(3). 237–264.
Levelt, W., A. Roelofs & A. Meyer. 1999. A theory of lexical access in speech production. *Behavioral and Brain Sciences* 22. 1–75.
McQueen, J. M. & A. S. Meyer. 2019. Key issues and future directions: Towards a comprehensive cognitive architecture for language use. In P. Hagoort (ed.), *Human Language: From Genes and Brain to Behavior*, 85–96. Cambridge, MA: MIT Press.
Menn, L. 1978. Phonological units in beginning speech. In A. Bell & J. Hooper (eds.), *Syllables and Segments*, 157–172. Amsterdam: North Holland.
Rose, Y. & B. MacWhinney. 2014. The PhonBank Project: Data and software-assisted methods for the study of phonology and phonological development. In J. Durand, U. Gut & G. Kristoffersen (eds.), *The Oxford Handbook of Corpus Phonology*, 380–401. Oxford: Oxford University Press.
Shaw, J., A. Gafos, P. Hoole & C. Zeroual. 2011. Dynamic invariance in the phonetic expression of syllable structure: A case study of Moroccan Arabic consonant clusters. *Phonology* 28(3). 455–490.
Slobin, D. 1978. A case study of early language awareness. In A. Sinclair, R. J. Jarvella & W. J. M. Levelt (eds.), *The Child's Conception of Language*, 45–54. (Springer Series in Language and Communication 2). Berlin/Heidelberg: Springer. DOI: 10.1007/978-3-642-67155-5_3.
Swingley, D. 2003. Phonetic detail in the developing lexicon. *Language and Speech* 3. 265–294.
Swingley, D. 2005. 11-month-olds' knowledge of how familiar words sound. *Developmental Science* 8(5). 432–443.
Tsuboi, N., W. Francis & J. Jameson. 2020. How word comprehension exposures facilitate later spoken production: Implications for lexical processing and repetition priming. *Memory* 29(1). 39–58. DOI: 10.1080/09658211.2020.1845740.
Wheeldon, L. & S. Monsell. 1992. The locus of repetition priming of spoken word production. *Quarterly Journal of Experimental Psychology* 44A(4). 723–761. DOI: https://doi.org/10.1080/14640749208401307.
White, K., J. Morgan & L. Wier. 2005. When is a *dar* a car? Effects of mispronunciation and referential context on sound-meaning mappings. In A. Brogos, R. Clark-Cotton & S. Ha (eds.), *Proceedings of the 29th Annual Boston University Conference on Language Development*, 651–662. Somerville, MA: Cascadilla Press.
Wijnen, F. 1990. The development of sentence planning. *Journal of Child Language* 17. 651–676.
Wijnen, F. 1992. Incidental word and sound errors in young speakers. *Journal of Memory and Language* 31. 734–755.