BUCLD 38 Proceedings
To be published in 2014 by Cascadilla Press
Rights forms signed by all authors

Two-year-olds Distinguish Snakes from Nakes but not Trains from Tains

Margarita Gulian[1], Caroline Junge[2] and Clara Levelt[1 3]

1. Introduction

When two-year-olds start producing words with onset clusters, they typically reduce the word by omitting one of the cluster consonants. However, as careful analyses of the produced tokens showed (Gulian & Levelt, 2009, 2011), at a certain developmental stage we find acoustic traces of the omitted consonant, unperceivable to the adult ear. This suggests that in the lexical representation, both consonants are present. Target clusters are, at this point, thus reduced at lower levels of the speech production mechanism. The evidence for this assumption so far only comes from analyses of child productions. The amount of detail in the lexical representations of clusters of two-year-olds can also be deduced from their perception of clusters. So far, only one study exists on cluster perception, which examined whether 6-9-month-olds are sensitive to phonotactic regularities in cluster consonants, but these concerned pseudo-words (Archer & Curtin, 2011), not words stored in the mental lexicon. Here, we report on a preferential looking experiment, which investigated whether two-year-olds distinguish between correct and reduced clusters of known words. We tested 24-month-olds, who were considered to have a vocabulary large enough to contain a variety of words with clusters in the onset, yet still reduce them in production.

To examine how detailed toddlers store their onset clusters in their mental representations, we carried out a preferential looking experiment ('PLP': Golinkoff, Hirsh-Pasek, Cauley and Gordon, 1987; for a recent review see Golinkoff, Ma, Song, and Hirsh-Pasek, 2013). Swingley and Aslin (2000) modified this paradigm to examine how detailed 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

[1] Leiden University; corresponding author: c.c.levelt@hum.leidenuniv.nl
[2] University of Amsterdam
[3] The authors would like to thank the babies who participated in this study and their parents, and Monique Bisschop for running the experiments and coding the videos. This work was supported by a VIDI grant (276-75-006) from the Netherlands Organisation for Scientific Research (NWO), awarded to Clara Levelt.

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 correctly produced than when it is misproduced.

Most studies comparing infants' performance for 'correct pronunciations' versus 'mispronunciations' provide evidence 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; Bailey & Plunkett, 2002; Mani, Mills & Plunkett, 2012; Mani & Plunkett, 2007; Swingley, 2003, 2005, 2009; Swingley & Aslin, 2000; White & Morgan, 2008; but see Mills et al., 2004, and Altvater-Mackensen, van der Feest & Fikkert, 2013). Infants notice mispronuncations both for well-known words and for recently-learned words (Bailey & Plunkett, 2002; White & Morgan, 2008), which suggests that lexical representation are phonetically detailed from an early age on (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 (Altvater-Mackensen et al., 2013), and on the overlap in phonological features between target and the substituted phoneme (White & Morgan, 2008).

We created 'mispronunciations' not by substituting one phoneme for another, but by reducing the target onset cluster to a single consonant, according to the predominant reduction patterns in child productions. Which consonant is omitted depends on the identity of the cluster. For consonants containing a liquid as the second consonant (/Cliq/ clusters), it is the second consonant that is omitted, whereas for clusters starting with an /s/ (/sC/ clusters) it is usually the first consonant, /s/, that is omitted (Fikkert, 1994; Jongstra, 2003). In the preferential looking experiment we therefore compared the perception of reduced vs. correct /Cliq/ clusters and reduced vs. correct /sC/ clusters. If two-year-olds have stored clusters correctly, we expect them to be sensitive to correct vs. reduced productions of these clusters.

2. Method

2.1 Participants

Data from 40 monolingual Dutch children (mean age: 24;06, range: 23;16 and 24;21; 20 girls) 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.

2.2 Stimuli

For the perception experiment we selected 27 words (11sC- cluster words; 11 Cliq-cluster words; six filler words that served as distractor words at test) that most two-year-olds would know (Bacchini, Boland, Hulsbeek, Pot & Smits, 2005). 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.

Auditory stimuli accompanying each picture were recorded in a soundproof booth, with a sample rate of 44.1 kHz. A native female 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 preceding 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 he!, 'Look, a [shoe]! Isn't it pretty?'. Target words are either a /Cliq/ word like bril 'glasses' or bloem 'flower', or a /sC/ word like schoen 'shoe'. Words were either correctly produced (CC) or reduced (RC). To illustrate, a /Cliq/ -word like bril was correctly produced as [bʀɪl] and reduced to bil [bɪl]; and an /sC/-word like schoen was correctly produced as [sχun] and reduced to choen [χun]. The mean duration of all correctly pronounced target words was 770 milliseconds (800ms for /sC/ words; 730ms for /Cliq/ words), while for all mispronounced target words it was 660 milliseconds (680ms for reduced /sC/ words; 640ms for reduced /Cliq/ words).

2.3 Procedure

The entire experiment took place in a 2m × 2m soundproof booth. During the experiment, infants sat on their caregiver's lap at a distance of 90 cm from the screen. One camera, mounted directly under the screen recorded the infant's eye movements. Caregivers wore headphones and listened to a mix of music and backward speech. 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 experiment consisted of a familiarization phase, followed by a test phase. The function of the familiarization phase was to make sure that children recognized the nouns that were presented in the experimental phase. In the familiarization phase, all 27 objects were presented in isolation: the picture of the target word was slowly moving up and down for four seconds, while the target word was named correctly at two seconds. The test phase consisted of 25 trials (22 test trials; 3 filler trials), in which toddlers saw two objects side by side moving slowly up and down while the auditory stimulus was presented, naming only one of the objects. Each experimental trial lasted for eight seconds. The target word was 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 Appendix 2). We controlled for category effects; pairings either consisted of two animate objects or two inanimate objects. Each pairing was presented twice, with both objects occurring once as the target.

At test, infants heard only one version of each target word: either produced correctly or reduced. 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. This entailed four versions of the experiment: each version was presented to 10 subjects. In each version, subjects were presented with a relatively equal ratio of correct and reduced pronunciations of target words, for both types of consonant-clusters. In total the experiment contained four experimental conditions: CC-Cliq, CC-sC, RC-Cliq, RC-sC. The different conditions of the target words are schematized in Table 2 below. All infants saw all conditions, with 5 or 6 trials per condition.

Table 2: Conditions of the experiment (example trial 'bril-schep' glasses-shovel)

trial type	cluster type	experimental conditions	example
correct cluster	/Cliq/ words	CCliq	bril [bʀɪʟ]
	/sC/ words	CsC	schoen [sχun]
reduced cluster	/Cliq/ words	RCliq	bil [bɪʟ]
	/sC/ words	RsC	chep [χɛp]

Trials were presented in a semi-randomized way: Two trials of the same condition or with the same pairing never followed one another. See Appendix 2 for the trial-overview. In total, the experiment lasted eight minutes.

2.4 Scoring and analyses

The looking behavior was coded off-line by trained scorers with Elan (EUDICO Linguistic Annotator) 3.6. 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 - 2000 ms; and the post-naming phase from 360 ms after the onset of the target word up to 2000 ms after the onset of the target word: 2360 – 4000 ms. The delay of 360 ms after the word onset is considered to be the time that infants need to initiate an eye movement in response to speech (Swingley & Aslin, 2000). To ensure that naming effects truly reflected the effect of naming, we only included trials in which both target and distracter were fixated in the pre-naming phase.

We used two measures: 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 for a phase by the total time spent looking at either the target or the distractor in this phase (Swingley & Aslin, 2000). Latency longest look is the difference between infant's longest look at the target and that at the distractor within one phase (cf. Mani & Plunkett, 2007). The effect of naming on a 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 infants fixated the target relatively more after than before it was named. If infants have stored their onset consonant clusters in detail, naming effects should be larger for correctly-produced consonant clusters than for reduced clusters. In our statistical analysis we use these difference measures for both the PTL and the LLK measures, separately for each of our four conditions (CC-sC; CC-Cliq; RC-sC; RC-Cliq).

3. Results

3.1 PTL measure

A repeated measures analysis of variance test (ANOVA) was performed, with pronunciation condition (correct cluster vs. reduced cluster) and cluster type (liquid cluster vs. /s/ cluster) as factors. This revealed that the factor of cluster type had a significant effect ($F [1,39] = 4.009$, $p = .05$): participants looked longer at sC-clusters than at Cliq-clusters. However, they did not differentiate between fully produced or reduced clusters (pronunciation: $p \geq .12$; interaction pronunciation x Cluster type $p > .11$).

Figure 1A plots the mean values for the difference PTL measure for all four conditions: all have significantly increases in looking times at the target word (i.e. compared to 0; $p \leq .034$). Exploratory T-tests revealed that when words are correctly produced, there is a significantly larger naming effect for /sC/ words compared to /Cliq/ words ($t(39) = -2.63$; $p = .012$). For reduced words, cluster-type did not affect the looking behavior ($p>.9$).

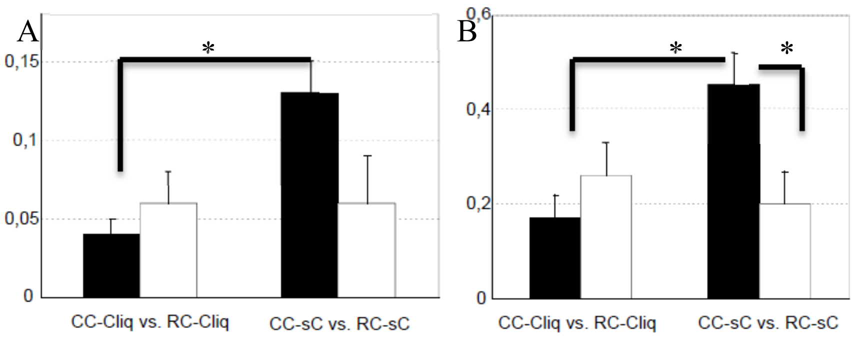

Figure 1: Mean increases in PTL on the left (A), and mean increases for LLK on the right (B), for the four conditions.

3.2 LLK measure

The results were corroborated in analyses with the LLK measure: there were no main effects of cluster type and pronunciation condition (all $p \geq .099$). However, there was a significant interaction between cluster type and pronunciation condition ($F [1,39] = 6.51$, $p = .015$). With exploratory T-tests, the difference LLK measure turned out to be significantly larger for correctly-produced /sC/ words compared to when these words were reduced ($t(39) = 2.28$; $p = .028$); this pattern was not observed for /Cliq/ words ($p \geq .35$). When contrasting the correctly produced conditions, we observed again that there were larger naming effects for the /sC/-clusters ($t(39) = -3.19$; $p = .003$). Figure 1B plots the mean increases of the four conditions for the LLK measure; again, all conditions elicited significant naming effects ($p \leq .012$); with largest increases for the correctly produced /sC/-clusters.

4. Discussion

We 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 found for correctly produced /sC/ clusters, and it differed significantly from the effect for the incorrect pronunciation. Given that most children of this age do not yet produce /sC/-clusters, their lexical representations appear to contain more detail than their own productions yet reveal. This finding is in line with our production study where we found compensatory lengthening in "numeral + /sC/ word" phrases, when /s/ was omitted (Gulian and Levelt, 2011). Our perception experiment revealed no significant difference for the correct vs incorrect /Cliq/ cluster targets.

It appears to be most obvious to relate the difference between cluster types to the fact that the incorrect pronunciation of /sC/ clusters involves the omission of the initial consonant, while it involves the second consonant in /Cliq/ clusters. Since the first consonant is very important for lexical access, it is likely that omission of this consonant will negatively influence lexical access.

Another possible explanation lies in the type of consonant that is missing. Even though /s/ is less sonorous than liquids according to the sonority scale (Selkirk, 1984), it is different in clusters, since, in the case of /s/ as C_1 and the liquid as C_2, the /s/ is acoustically more salient than the liquid. Furthermore, it has been argued that /s/ in onset clusters has a different status as compared to other consonants in C_1 position in the cluster: it should be viewed as an appendix to the syllable rather than as part of it (Fikkert, 1994). So, when children reduce /sC/ and /Cliq/ clusters in production, the source of these reductions could be of a different nature. While /sC/ cluster reduction might illustrate a phonological encoding problem (i.e. appendix syllabification; but all phonemes are stored), the source of /Cliq/ reduction could be the (weak) lexical representation of the /+liq/ part.

Recall however that careful analyses of production data reveal small traces of liquids in second positions (Gulian & Levelt, 2009). This points to the possibility that words with /Cliq/ clusters are stored as C_1C_2V form in the mental lexicon but that the C_2 is less specified than C_1. If children have a more detailed representation of the word edges than of the C_2, a reduced form like 'tain' will be equally 'distant' from the segmental representation as 'train' (with a full /r/) and both forms permit access to the concept of a train. Nevertheless, based on the perception study presented above, it is difficult to tear apart whether children for /Cliq/ clusters have not stored the second consonant at all, or whether this consonant is less-specified. Clearly, more (cross-linguistic) research is needed to distinguish between these two accounts.

We further observed that children looked longer at the target picture when it was a correctly produced /sC/ word than when it was a correctly produced /Cliq/ word (mean difference 230 ms). However, children did not exhibit such a preference before a target was named: The LLK did not appear to differ significantly in the pre-naming-phase of these two conditions ($p \geq .36$), while it did differ significantly in the post-naming phase ($t(39) = -3.03$, $p = .004$; similar results for the PTL-measure). Duration (alone) cannot explain this pattern: the mean duration of /sC/ words was only 55ms longer than that of /Cliq/ words, which is smaller than the 230 ms difference in looking times in the post-naming phase. We therefore propose that the data show a "learning effect" for /sC/ clusters. These clusters are usually acquired later than /Cliq/ clusters (Fikkert 1994). The learning effect results from the learners' comparison of their own reduced form, e.g. [tul], which is probably stored as such (/tul/) to the perceived correct form /stul/. The presence of /s/ in the correct form is noticed and attracts attention, which leads to longer looking times. It is interesting to note that preliminary production data of the same participants in our experiment indeed appear to show that the strongest naming effect for correct sC clusters is found for those who reduced /sC/ clusters in their productions. No such effect is found for the /Cliq/ clusters. The stored lexical representation for words with these clusters is probably already correct, and the reduced form is simply accepted as close enough to recognize the word to the same extent. Of course, such a learning effect does not preclude the possibility that children noticed a mispronunciation effect at the same time: looking times to correctly produced /sC/ words were both longer for correctly produced /Cliq/ words and for reduced /sC/ words. Rather, an account in terms of a learning effect presents a different approach to our data.

5. Conclusion

In this paper we examined the way in which 24-month-olds perceive different reduced onset clusters. Results from this experiment suggest an asymmetry: participants differentiated between correctly produced and reduced /sC/ cluster words, but not between correctly produced and reduced /Cliq/ words. That children noticed the omission of /s/ in /sC/ cluster words is interesting, because children this young generally have difficulty producing the fricative in /sC/ clusters (Fikkert, 1994; Jongstra, 2003; Smit, 1993). At the same time, our results can be interpreted as evidence that two-year-old Dutch children may have a more detailed representation of the first segment of a cluster word than of the consonant in second position. Either way, two-year-olds do not yet store all consonant clusters with the same detail.

We further suggest that the longer LLK at the target picture for correct /sC/ cluster words can be understood as a learning effect. The longer looking time is a sign of awareness of the discrepancy between the perceived correct form and their own reduced form, and therefore as a sign that learning of the correct form is taking place. Participants who have not acquired /sC/ clusters yet appear to be the ones who specifically exhibit this effect. The exact relation between production and perception in development needs to be established further in future work, since it appears to inform both the interpretation of perception experiments and the source of deviating productions.

Appendix 1

For each category of words, all the target words and their English translation in italics:

/Cliq/-words (n=12):	vlag *flag*; brood *bread*; broek *trousers*; trui *sweater*; bril *glasses*; blok *block*; trap *stairs*; klok *clock*; kraan *tap*; trein *train*; fles *bottle*; bloem *flower*.
/sC/-words (n=10):	schep *shovel*; schaar *scissors*; schoen *shoe*; step *step*; stoel *chair*; spin *spider*; slang *snake*; schaap *sheep*; slak *snail*; speen *pacifier*.
Control words (n=5):	baby *baby*; bal *ball* ; auto *car*; poes *cat*; mier *ant*.

Appendix 2

Table of test trials for List 1, with information about picture-combinations, target (annotated with Dutch ortography), and conditions (CC= correct cluster; RC= reduced Cluster; sC=clusters starting with /s/; Cliq= clusters with liquid in second position). List 2 had the same pairings (but in a different order), with the target counter-balanced with List 1: correctly produced targets of one object were replaced by reduced clusters of the other object in a pairing, and vice versa (when a word was paired with a simple onset-word, as in trial 4 and 20, counter-balancing between lists was RC vs. CC of the same cluster). List 3 and 4 were identical to List 1 and 2, respectively, but in reverse order.

Trial	Pictures left-right	Labels in Dutch	Target	condition
1	flower-step	bloem-step	step	CC-sC
2	flag-trousers	vlag-broek	vlag	CC-Cliq
3	bread-bottle	brood -flesje	bood	RC-Cliq
4	kitty-sheep	poes-schaap	schaap	CC-sC
5	spider-snake	spin-slang	lang	RC-sC
6	glasses-shovel	bril-schep	schep	CC-sC
7	car-baby	auto-baby	baby	filler
8	trousers-flag	broek-vlag	boek	RC-sC
9	sweater-pacifier	trui-speen	trui	CC-Cliq
10	shoe-tap	schoen-kraan	choen	RC-sC
11	snake-spider	slang-spin	spin	CC-sC
12	shovel-glasses	schep-bril	bil	RC-Cliq
13	block-scissors	blok-schaar	bok	RC-Cliq
14	chair-stairs	stoel-trap	trap	CC-Cliq
15	baby-car	baby-auto	paku	filler
16	clock-train	klok-trein	klok	CC-Cliq
17	ball-kitty	bal-poes	bal	filler
18	stairs-chair	trap-stoel	toel	RC-sC
19	tap-shoe	kraan-schoen	kraan	CC-Cliq
20	ant-snail	mier-slak	slak	CC-sC
21	train-clock	trein-klok	tein	RC-Cliq
22	pacifier-sweater	speen-trui	peen	RC-sC
23	bottle-bread	flesje-brood	fles	CC-Cliq
24	step-flower	step-bloem	boem	RC-Cliq
25	scissors-block	schaar-blok	schaar	CC-sC

References

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

Altvater-Mackensen, N., van der Feest, S. V., & Fikkert, P. (2013). Asymmetries in Early Word Recognition: The Case of Stops and Fricatives. *Language Learning and Development, (ahead-of-print)*, 1-30.

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

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

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

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

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

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

Gulian, M. & Levelt, C. (2011). Temporal measures of reduced /sC/-clusters in toddler speech: evidence for a detailed lexical specification, *Online Proceedings ICPhS XVII Hong Kong.*

Gulian, M. & Levelt, C. (2009). An acoustic analysis of child language productions with reduced clusters. *Online proceedings BUCLD 33.*

Jongstra, (2003). W. Variable and stable clusters: Variation in the realization of consonant clusters. *Canadian Journal of Linguistics, 48, 3/4*, 265-288.

Mani, N., Mills, D. L., & Plunkett, K. (2012). Vowels in early words: an event-related potentiaal study. *Developmental Science, 15,* 2-11.

Mani, N., & Plunkett, K. (2007). Phonological specificity of vowels and consonants in early lexical representations. *Journal of Memory and Language, 57,* 252-272.

Mills, D.L., Prat, C., Zangl, R., Stager, C.L., Neville, H.J., & Werker, J.F. (2004). Language Experience and the Organization of Brain Activity to Phonetically Similar Words: ERP Evidence from 14- and 20-Months-Olds. *Journal of Cognitive Neuroscience, 16,* 1452-1464.

Selkirk, E. (1984), On the major class features and syllable theory. *Language sound structure.* Eds. M. Aronoff and R. Oehrle. Cambridge, London: MIT Press, 107-136.

Smit, A. B. (1993). Phonologic error distribution in the Iowa-Nebraska Articulation Norms Project: Consonant Singletons. *Journal of Speech and Hearing Research, 36*, 5 533-547.

Swingley, D. (2003). Phonetic detail in the developing lexicon. *Language and Speech, 2003, 46 (2-3)*, 265-294.

Swingley, D. (2005). 11-month-olds' knowledge of how familiar words sound. *Developmental Science, 8,* 432-443.

Swingley, D. (2009). Onset and codas in 1.5-year-olds' word recognition. *Journal of memory and language, 60*, 252-269.

Swingley, D. & Aslin, R. N. (2000), Spoken word recognition and lexical representation in very young children, *Cognition, 76*, 147-166.
White, K. S. & Morgan, J. L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language, 59*, 114-132.