Acquiring voicing in Dutch

The role of function words*

Suzanne van der Feest
University of Nijmegen

1. Introduction

In order to build up a representation of the Dutch voicing contrast in the lexical representations of words, the learner must first identify the relevant phonological contrast in the language. In Dutch, the main acoustic cue for voicing in initial position is Voice Onset Time (VOT): Voiced stops (e.g. /b/, /d/) have a negative VOT value of around – 4 ms., while voiceless stops (e.g. /p/, /t/) have a VOT value between 0 and 25 ms. (Lisker & Abramson 1964, among others). Thus, children acquiring Dutch need to learn to perceive and produce this VOT contrast. This paper will discuss early production data from Dutch children aged 1;0–2;8. The production of word-initial voicing will be discussed in Section 2 and 3. In Section 4, we will explore the hypothesis that the acquisition of the word-initial /d/–/t/ contrast is influenced by variation in voicing of /d/-initial function words. We argue that in the early stages of acquisition, this variation is what leads to an apparent delay of the development of the voicing contrast in coronals as compared to labials, and to more variation in productions within /d/-initial word-types. The conclusions are summarized in Section 5.

2. Production of plosives in initial position

To investigate the production of voicing in initial position, longitudinal data from 11 Dutch children were examined. The data were taken from the CLPF database from CHILDES (Fikkert 1994, Levelt 1994). This database contains ~20,000 utterances of monolingual Dutch children between the ages of 1;0–2;8. All word-initial /b/, /d/, /p/ and /t/ target segments that children produced with

Linguistics in the Netherlands 2004 21 (2004), 47–57.
ISSN 0929–7332 / E-ISSN 1569–9919

correct place of articulation were coded for how voicing was realized. For instance, initial /b/'s were coded for whether the child produced them as /b/ (correct, e.g. *bus* (bus) (Tom, 1;5) or as /p/ (incorrect, e.g. *pus*, Noortje, 2;1). Similarly, voiceless segments such as /p/ were coded as /p/ (correct, *papa* (Catootje, 1;11) or as /b/ (incorrect, *bapa*, Robin, 1;8). Note that /k/ was not coded, because Dutch lacks a voicing contrast in velar stops. Also, fricatives were not examined because the voicing contrast in fricatives is disappearing in large areas of the Netherlands (Ernestus 2000, a.o.).

Figure 1 shows the error pattern collapsed across all children. The errors in initial /b/, /d/, /p/ and /t/ were collapsed, thus showing the mean error percentages in the production of voicing in labial and coronal stops. The main conclusion that can be drawn from these data is that children acquiring Dutch have not yet completely mastered the voicing contrast by the age of 2;6. This has also been shown by Beers (1995) and Kuijpers (1993).

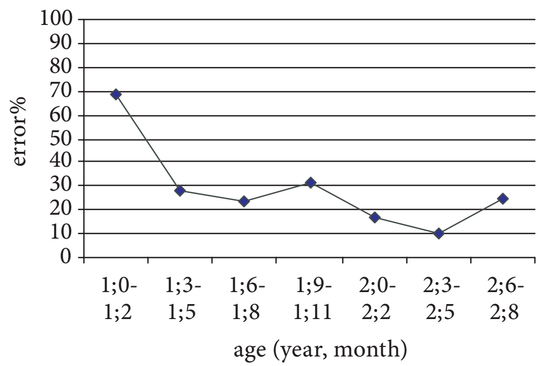

Figure 1. Overall error % in initial /b/, /p/, /d/ and /t/

In Figure 2, the errors have been split up according to place of articulation: this figure shows the errors made in coronal stops versus the errors made in labial stops. When we look at the group results, there is no main effect for place of articulation. While coronals appear to be produced less faithfully, there was no significant difference (*t-test*, $p=0.15$, two-tailed) between the errors in the productions of coronals (M = 24.3, SD = 9.04) as compared to labials (M = 21.5, SD = 6.3). Voiced and voiceless stops are not produced more or less faithful in labials than in coronals.

However, when the errors are split into voiced targets (/b/ and /d/) versus voiceless targets (/p/ and /t/), a clear effect is visible. In Figure 3 voiced targets produced incorrectly are coded as devoicing errors (e.g., /b/ and /d/ become /p/ and /t/, respectively) and voiceless targets produced incorrectly are coded as voicing errors (e.g., /p/ and /t/ become /b/ and /d/, respectively). We can now see that there were more devoicing errors (M = 32, SD = 11.6) than voicing

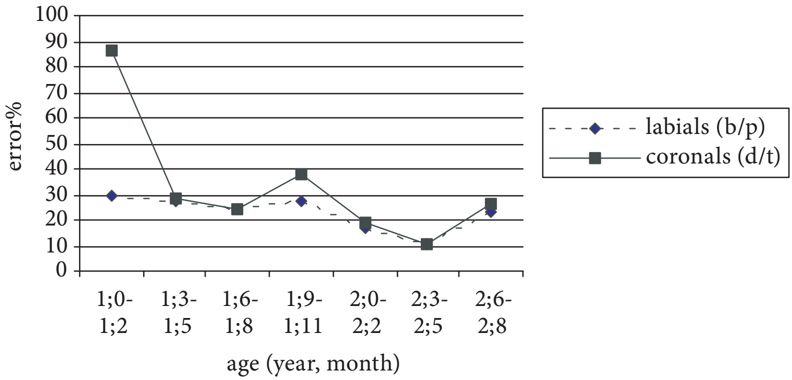

Figure 2. Direction of errors — place of articulation

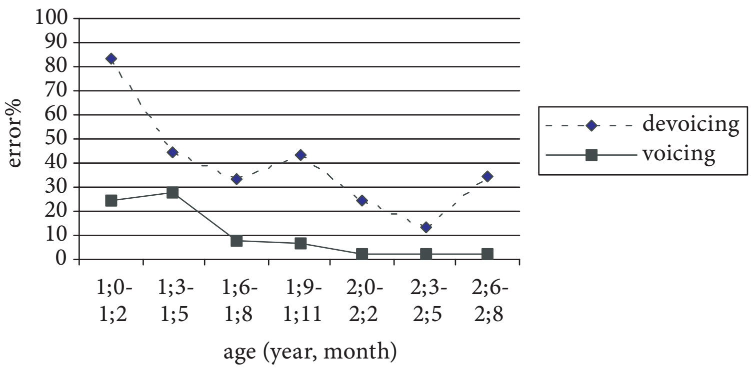

Figure 3. Direction of errors — voicing

errors (M = 8.17, SD = 10.08). This difference was significant (*t-test*, $p \leq 0.01$, two-tailed).

Interestingly, when we look at the distribution of voicing in initial position for the target-words that children attempt, it turns out that there are more voiced than voiceless initial targets overall (4871 voiced targets versus 2244 voiceless targets words). However, even though more target words have initial voiced stops, children are more accurate in producing voiceless stops (see Figure 3). This shows that the frequency of voicing in the targets that children attempt to produce does not reflect the order of acquisition. Therefore, frequency cannot account for the difference in acquisition between voiced and voiceless stops in Dutch.

On the basis of these Dutch data and based on acquisition data from German and English, Kager et al. (forthcoming) argue that the privative feature [voice] is the active and marked feature in Dutch. This means that voiceless segments are unmarked for voice. If children have not yet specified the voicing feature in the lexical representation of a word, they will tend to produce the unmarked feature value, resulting in voiceless productions of voiced stops. In

Section 4.2, it will be argued that this is what happens in productions of Dutch /d/-initial function words, which remain unspecified longer due to the variation children encounter in the input.

3. Acquisition of /d/ versus /b/

When the data were collapsed across children, no significant effect was found for how voicing was realized in different places of articulation. However, a significant effect is found when we look at the data from individual children. When comparing errors in the realization of voicing per segment, the patterns of four of the eleven children show more errors in coronal- than in labial-initial words. The data from these four children are shown in Figure 4. This figure shows the error percentages for /b/ versus /d/-initial words.

The children that show this pattern are the youngest children in the CLPF database (Tom 1;0–2;3, Eva 1;4–1;11, Jarmo 1;4–2;4 and Leonie, who was monitored for a short period between 1;9–1;11). The age difference with the other children in the database is minimal, but the other children appear to be in a later stage of the acquisition of the voicing contrast.[1]

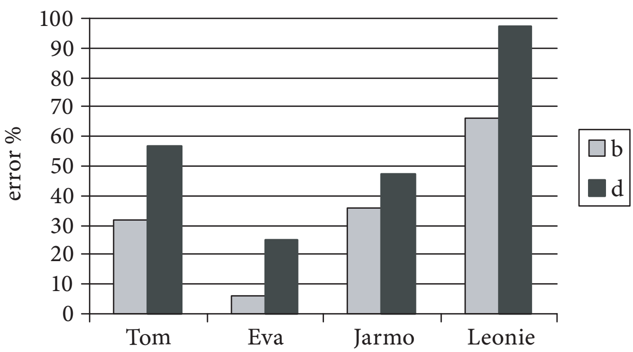

Figure 4. Graph with four children, /b/ versus /d/

The difference between the errors in /b/ versus /d/-initial words in this group of four children was significant ($\chi^2=23.49$, $p\leq0.001$, 980 tokens), as was the difference between coronals (/d/ plus /t/) and labials (/b/ plus /p/) ($\chi^2=24.85$, $p\leq0.001$, 1552 tokens). We also conducted χ^2 for individual children — the pattern held for each child. Thus, the youngest children show an effect for place of articulation, which was not shown for the group of children as a whole across all stages. We next examined whether the older children showed an effect for place of articulation when only the initial-voiced targets (i.e. /b/ and /d/ targets) were taken into account. We found that there was no difference between the number of errors made in /d/ as opposed to /b/-initial words

($\chi^2 = 1.20$, n.s., 3903 tokens). Moreover, the opposite pattern is never shown: no children in the database make significantly more errors in /b/-initial words.

This suggests that the /d/–/t/ contrast is acquired slightly later than the /b/–/p/ contrast. An explanation for this could be given by an articulatory account — because the closure of the mouth is complete when producing labials, it is easier to produce voicing in a labial (/b/) than in a coronal stop (/d/) (e.g. Van Alphen, 2004). However, a closer examination of the difference between content words and function words suggests that this explanation cannot account for the acquisition patterns discussed in this paper. Instead, we hypothesize that the difference between the error rates in labials versus coronals can be explained by a possible influence from the variation that children encounter in the input. This variation comes from the voicing in /d/-initial function words. We will now turn to a discussion of the role of function words in Dutch.

4. Function words

4.1 Content versus Function words

Many function words in Dutch are /d/-initial, such as *die* (that), *deze* (this), *dit* (this), *dat* (that), *daar* (there) and *de* (the). These function words are very frequent in Dutch, and function words comprise a large proportion of the /d/-initial target words children attempt — 23.9% of all examined target words were function words and 39% (1701 tokens) of all /d/-initial target words consisted of function words. Dutch has a few /t/, /b/ and /p/-initial function words (for example *bij* (with) and *te* (particle 'to'), but these words were not or very rarely produced by the children in the database. The majority of the produced function words (99%) were /d/-initial. In general, function words are phonologically less salient than content words (function words tend to have for example shorter vowel duration, weaker amplitude and simplified syllable structure compared to content words) (Shi et al. 1999).

In utterance-medial position there is a lot of variation in how voicing is realized in these Dutch function words (Ernestus 2000, among others). For example, '*Wat is dat?*' ('What is that') is often produced as '*Wat is tat*', with a devoiced /d/ in the function word '*dat*'.[2] This variation is allowed in utterance-medial position in standard Dutch. However, 66% of all plosive initial words in the database were produced in utterance-initial position (4699 out of 7115 tokens). In this position, variation in the voicing of the initial segment is not allowed in standard Dutch. This means that the children in the database were not merely devoicing segments where this is optional in Dutch, but also in utterance-initial position where this does not occur in adult speech. Also, the data of the four children who made more voicing errors in /d/-initial than in

/b/-initial words cannot be explained by a higher percentage of voiced stops occurring in utterance-medial position. On the contrary, these children produced 82% of their target words in utterance-initial position: thus, they devoice even more often in a position where this is not allowed.

Our next step was to determine whether the errors children made in the voicing of word initial /d/ came primarily from function words. In order to do this, we looked at children's production of voicing in content words only. When we look at the error percentages of just /d/-initial content words, the patterns of Tom, Eva, Jarmo and Leonie, match up more with the patterns of the other seven children's production of /d/-initial words (collapsed across function words and content words, as well as in content words only)

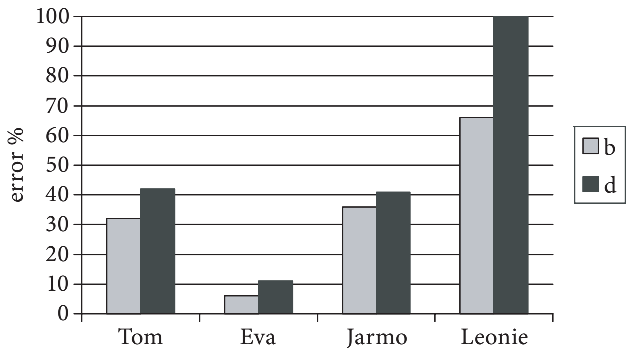

Figure 5. /b/ versus /d/, content words only

There is no longer a significant difference between their errors in /b/ and /d/-words collapsed across these four children ($\chi^2 = 5.08$, n.s., 624 tokens). At the level of individual children, only Leonie still made significantly more errors in /d/-initial content words than in /b/-initial content words ($\chi^2 = 9.42$, $p \leq 0.01$, 78 tokens). Thus, only for Leonie the claim could hold that voicing in coronals is harder to produce than voicing in labials. Note that Leonie was monitored for a short period of time, therefore providing less data than the other children. The productions of all other children in the database do not seem to be in line with an articulatory account.

These data show that the delay of the /d/–/t/ contrast versus the /b/–/p/ contrast in the children's earliest productions can be attributed to the errors in /d/-initial function words. In the next section we will explore the hypothesis that children leave the voicing specification in function words underspecified in the earliest stages of acquisition. These representations remain underspecified longer because children have not yet learned in which positions the variation in voicing is allowed: /d/-initial function words are thus produced mainly voiceless (the unmarked value). This is what results in the delay of the acquisition of /d/ versus /b/. If we assume that /d/ is underspecified in function words, the question is whether /d/ is only underspecified in function words or also in content words. This will be examined in the remainder of this paragraph. Section 4.2 will discuss previous research on the recognition of function words — in Section 4.3 we will examine the variation that occurs within the types that children produce. We will investigate whether the variation in /d/-initial function words also leads to more variation within the productions of /d/-initial content word-types.

4.2 The role of function words

The function versus content category difference is universally present in human languages. Previous research has shown that newborn infants (1 to 3 day-olds) are already able to discriminate lists of English function words from lists of English content words. (Shi et al. 1999) All tested infants were able to do this — even when the first language of their mother was not English. Thus, infants seem to have the ability to pick up on the typical prosodic features of content words versus those of function words. Six month-olds are still able to discriminate between the two categories, and even prefer to listen to lists of content words over lists of function words (Shi & Werker 2001). Further research has shown that 13 month-olds (but not 8 month-olds) can discriminate between real function words (e.g. *the*, *his*) and minimally different nonsense function words (e.g. *kuh*, *ris*) (Shi et al. 2003). These findings might suggest that 13 month-old children already have a detailed lexical representation of function words, and that infants from birth are already able to pick up on the differences between content and function words. This would not be in line with our hypothesis that children maintain an underspecified representation of voicing in function words. However, the following points should be considered.

First of all, studies on the recognition of function words have not tested recognition of real versus nonsense function words when only the laryngeal feature was changed in the non-word (e.g. the Dutch function word *daar* (there) changed into *taar*). It is thus still possible that children do not have a detailed stored representation of the voicing features of function words. Furthermore, Dutch children encounter variation in the realization of voicing in function words in the input. This variation is restricted to certain contexts, but these contexts do include positions where variation in content words is not allowed. As was mentioned in Section 4.1, '*Wat is dat*' (what is that) can be pronounced as '*Wat is tat*' with a devoiced /d/. This however is impossible in a phrase like: '*Wat is duiken?*' (what is diving): *'W*at is tuiken*' is not allowed. Thus, our claim is that despite the fact that young infants are able to differentiate function words from content words and function words from nonsense function words, Dutch children at the earliest stages of production do not yet have a detailed stored representation of voicing features in function words.

Children need to acquire in which contexts the variation in voicing is allowed — only when they have learned that the variation is not random, can they specify the voicing value of the initial segment in their lexical representations of function words.

4.3 Variation within types

If children have underspecified voice representations in /d/-initial function words, what does this mean for the acquisition of the /d/–/t/ contrast in general? In Sections 2 and 3 we showed that children did not produce significantly more errors in /d/- than in /b/-initial content words. Our next question is whether the variation in /d/-initial function words only leads to more devoicing of these function words than is allowed, or whether there is also an effect on the productions of /d/-initial content words. If such an effect would occur, then children initially will allow for more variation in /d/-initial words than in /b/-initial words. In order to test this, we analyzed the variation within word-types. For each plosive-initial type (all /d/, /t/, /b/ and /p/ types), we checked whether the voicing in this word was produced consistently or with variation, i.e. produced either correct or incorrect by the same child at the same age. This would mean that for example the /p/ in *paard* (horse) was sometimes produced correctly (as *paard*) and sometimes incorrectly (as *baard*). Figure 6 shows the percentage of word-types that were produced with variation by the same child. The data of all eleven children have been collapsed in this graph. The error rates in /d/-initial word-types are split up into function words versus content words.

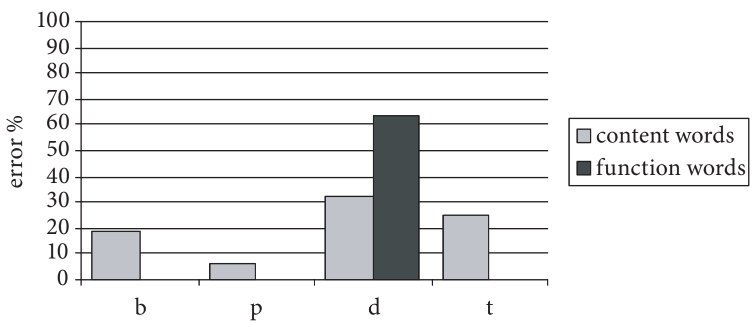

Figure 6. Variation within word-types — all eleven children

When looking at all types (function words and content words together), there was significantly more variation within coronal-initial types than within labial-initial types. This held for the four youngest children ($\chi^2=13.45$, $p\leq 0.001$, 648 types) as well as for the whole group ($\chi^2=48.26$, $p\leq 0.001$, 3005 types). This place effect can be ascribed to the fact that the variation within types is larger in function words than in content words. (The difference between the variation in content words versus function words in /d/ was significant ($\chi^2 = 90.6$, $p \leq 0.001$, 3021 types).) When looking at content words only, there was no longer a significant effect for place of articulation, neither for the group as a whole (($\chi^2 = 1.03$, n.s., 2686), nor for the youngest age group (($\chi^2 = 4.45$, n.s. on a $p \leq 0.01$ level, 742 types).

However, when we looked at /d/-initial words only, we did find a difference between the four youngest children and the seven older children in the database. For the older children, the variation within types of /d/-initial function words was significantly greater than in /d/-initial content words ($\chi^2 = 18.23$, $p \leq 0.001$, 674 types). This effect was also significant for the group of children as a whole ($\chi^2 = 16.72$, $p \leq 0.001$, 887 types). Thus, as would be expected on the basis of the input, children show more within-type variation in function words than in content words. This turned out not to be the case for the youngest children. In /d/-initial words, Tom, Jarmo, Eva and Leonie show just as much within-type variation in content words as in function words: the difference between them was not significant ($\chi^2 = 1.81$, n.s. 211 types). They show more variation in /d/-initial content words than the older children ($\chi^2 = 20.13$, $p \leq 0.001$ 894 types). This suggests that the variation in the /d/-initial function words carries over to the /d/-initial content words in the productions of these youngest children. An articulatory account predicting that voicing in /d/ is harder than voicing in /b/ would not account for these data: when voicing in /d/ is acquired later, we would expect more errors in /d/-initial words overall. However, the only effect we found was not for the overall errors in /d/, but for the variation within /d/-initial types compared to /b/-initial types. Thus, we claim that the variation in the function words in the input is taken into account when children are acquiring the /d/–/t/ contrast, resulting in more within-type variation in /d/-initial content words in the initial stages of acquisition.

5. Conclusions and discussion

Function words clearly do not facilitate the acquisition process of initial voiced coronal stops in Dutch. They form a highly frequent category in Dutch, and they are essential for syntactic bootstrapping in the early stages of acquisition. Therefore, infants' ability to distinguish content words from function words on the basis of their prosodic characteristics is crucial. The prosodic cues that make them different, such as the ones mentioned in Section 4.1, are therefore crucially present. But this does not mean that function words also provide clear evidence for the phonology of a language — in this case, for the voicing system of Dutch. In fact, assuming that children take all the acoustic input into account when acquiring phonological contrasts, /d/-initial Dutch function words make this process harder. The variation children encounter in these function words prevents them from specifying the voicing value in their early lexical representations of these words. This results in an apparent delay in the acquisition of the initial /d/–/t/ contrast compared to the initial /b/–/p/ contrast. We have tried to show that this cannot be explained by an articulatory account stating that voicing in /d/ is simply harder to produce. We argued that the difference can be ascribed to the influence of the function words. The laryngeal features in the representations of function words will remain unspecified longer than those in content words — where the child does not encounter variation. Thus, there is a carry-over effect to contexts (utterance-initial position) where this variation is not allowed. Initially, there is also a carry-over effect to /d/-initial content words, where the younger children allowed for more variation within types than the older children.

In conclusion, patterns in initial voicing in the productions of these Dutch children can give us an insight into what their early lexical representations may look like. Perception experiments with children in the same age ranges might provide further evidence for our claims on voicing in early lexical representations, and the role of the function words.

Notes

* The author would like to thank the members of the NWO project group 'Development of Phonological Representations for Perception and Production' (grant 360–70–100) and the audience at the TIN-dag 2004 for their comments.

1. This does not count for Noortje who was late in her overall phonological development (Fikkert 1994). She made more than 75% errors in both /b/ and /d/ (The difference between /b/ and /d/ was not significant ($\chi^2 = 0.0003$, n.s., 436 tokens).

2. Dutch is an interesting language, since both progressive and regressive voicing assimilation occur. Ernestus (2000) claims that the voicing of the /d/-initial function word in utterance position determines the voicing value of the preceding segment. Thus, when the function word is realized as voiced, this leads to regressive voicing assimilation and 'Wat is dat' will produced as '*Wat iz dat*' with a voiced fricative. When the function word is produced voiceless, which is optional, the phrase will be produced as '*Wat is tat*' with a voiceless fricative.

References

Alphen, P.M. van (2004). *Perceptual relevance of voicing in Dutch.* PhD dissertation, University of Nijmegen. MPI series in Psycholinguistics, Wageningen.

Beers, M. (1995). *The phonology of normally developing and language-impaired children.* PhD Dissertation, University of Amsterdam.

Ernestus, M. (2000). *Voice Assimilation and Segment Reduction in Casual Dutch. A Corpus-based study of the phonology-phonetics interface.* PhD Dissertation, Vrije Universiteit Amsterdam. LOT, Utrecht.

Fikkert, P. (1994). *On the acquisition of prosodic structure.* PhD dissertation, Leiden University. HIL dissertations, The Hague.

Kager, R., S.V.H. van der Feest, P. Fikkert, A. Kerkhoff & T.S. Zamuner (to appear). 'Representations of [voice]: Evidence from Acquisition.' In: J. van de Weijer & E.J. van der Torre (eds.) *Voicing in Dutch.*

Kuijpers, C. (1993). *Temporal coordination in speech development: a study of voicing contrast and assimilation in voicing.* PhD dissertation, University of Amsterdam.

Levelt, C.C. (1994). *On the acquisition of place.* PhD dissertation, Leiden University. HIL dissertations, The Hague.

Lisker, L. & A.S. Abramson (1964). 'A cross-language study of voicing in initial stops: acoustical measurements.' In: *Word* 20, 384–422.

Shi, R. & J. Werker (2001). 'Six-month-old infants' preference for lexical words.' In: *Psychological Science* 12, 70–75.

Shi, R., J. Werker & A. Cutler (2003). 'Function words in early speech perception.' In: *Proceedings of the 15th International Conference of Phonological Sciences.*

Shi, R., J. Werker & J. Morgan (1999). 'Newborn infants' sensitivity to perceptual cues to lexical and grammatical words.' In: *Cognition* 72, B11-B21.

Swingley, D. (2003). 'Phonetic Detail in the Developing Lexicon.' In: *Language & Speech* 46, 265–294.

Swingley, D. & R.N. Aslin (2000). 'Spoken word recognition and lexical representations in very young children.' In: *Cognition* 76, 147–166.