Fifteen-month old infants' sensitivity to vowels' first and second formants in novel word learning

Titia Benders[1]
Thesis Research Master in Linguistics
University of Amsterdam

Supervisors:
Paola Escudero (University of Amsterdam), first reader
Suzanne Curtin (University of Calgary), second reader
Daniel Swingley (University of Pennsylvania)

---

[1] I would like to thank my three supervisors for their help and advice before and throughout the experiment. A special thank-you is for Suzanne Curtin, who took the risk of letting me stay in her lab for six months. Within Suzanne's Speech Development lab, I am very much indebted to Stephanie Archer for lending me her voice, face and time for the stimuli and to Natasha Nickel for helping me out in all aspects of the testing process. Jane Park in Daniel Swingley's Infant Language Lab was very helpful in finding pictures of the stimuli. Thanks to Paola and Suzanne for their quick comments on an earlier version.
I was funded by a Stunt-grant form the University of Amsterdam, a travel grant from the Amsterdam University Association, and study grants from the Dr. Hendrik Muller's Vaderlandsch Fonds, the Fundatie van de Vrijvrouwe van Renswoude te 's Gravenhage and the Schuurman Schimmel-van Outeren Stichting. I also received support from NOH grant no. 1-R01-HD49681 to Daniel Swingley, and NWO VI grant no. 016.024.018 to Paul Boersma.
For correspondence: Regular mail: Titia Benders, Amsterdam Center for Language and Communication, University of Amsterdam, Spuistraat 210, 1012 VT Amsterdam, The Netherlands. E-mail: titia.benders@uva.nl. Phone: 0031-(0)20-525-8148.

1. INTRODUCTION

Newborn infants already have abilities that could help them in their later language acquisition, such as discriminating between languages from different rhythmic patterns (Nazzi, Bertoncini, & Mehler, 1998) but they do not yet understand any words. During the first half year of their lives, infants begin to learn the connections between highly frequent words and highly familiar referents, such as the words for their parents (Tincoff & Jusczyk, 1999) and this develops towards the recognition of some words for familiar objects at 13 months (Thomas, Campos, Shucard, Ramsay, & Shucard, 1981). Word learning really kicks off in the second year of life, with the so-called 'vocabulary spurt' or 'naming explosion' (Benedict, 1979). The vocabulary spurt is mainly observed for children's productive vocabulary size after the age of 18 months. Researchers of child language acquisition have asked many questions about children's knowledge of these early words, acquired prior to and during the vocabulary spurt. This thesis stands in the tradition of questions about infants' knowledge of the forms of these early words.

Although there are many different theories about what it takes to learn a word form and then recognize it upon hearing auditory input (e.g. (Goldinger, 1996) for the stance that exemplar-like episodic traces are stored and used in word recognition, and (Lahiri & Reetz, 1999) for the position that only abstract phonological features play a role), the three general steps given in Table 1 figure in most theories. In this thesis I am mainly concerned with step 3 of word learning and word recognition: which information is stored in the word representations and which encoded properties of the speech signal matter in accessing these representations?

step	Learning	Recognizing
1	Attending to the speech signal	
2	Encoding the properties of the speech signal	
3	Storing the properties of the speech signal in a representation	Matching the encoded properties on the stored representation.

Table 1 Three steps that play a role in word learning and subsequent recognition of these words.

In the remainder of this introduction, I will first discuss the two main paradigms that are being used to investigate young infants' word recognition and word learning abilities, the switch paradigm (Werker, Cohen, Lloyd, Casasola, & Stager, 1998) and the two-alternative forced choice paradigm (Golinkoff, Hirshpasek, Cauley, & Gordon, 1987) with mispronunciations (Swingley & Aslin, 2000), and potential causes for the differences in outcomes between the two paradigms. I will specifically focus on what can be concluded about infants' knowledge of word forms from the absence of an effect in either paradigm.

In the second part of this introduction I will turn to research within both paradigms that has focussed on specific phonetic details in infant word recognition. I first discuss studies that have shown that for consonants, there are asymmetries in infant word recognition for consonants (e.g. (van der Feest, 2007)). Then I will discuss two recent studies that have investigated the effect of vowels' first formant (F1), corresponding to phonological vowel height, and second formant (F2), corresponding to phonological vowel backness, on infants' word recognition (Curtin, Fennell, & Escudero, In press; Mani, Coleman, & Plunkett, 2008). These studies yielded different results regarding infants' sensitivity to F2. This thesis aims at gaining more insight in the cause of these differential findings. Because perceptual asymmetries are often observed in vowel perception (cf. (Polka & Bohn, 2003) for a review), the secondary goal of this thesis is to explore asymmetries for vowels in infant word recognition.

1.1 METHODS USED IN INFANT WORD RECOGNITON RESEARCH

1.1.a The switch paradigm

The first method I discuss here is the switch paradigm, introduced by (Werker et al., 1998). The switch paradigm is a habituation paradigm. In a habituation experiment, participants are presented with the same stimulus over and over again until their attention, for infants often measured in looking time to a visual stimulus, decreases to a pre-set criterion, for instance 50% of the attention to the first block of trials. When the criterion is reached, a test stimulus is presented, which is slightly different from the habituation stimulus. If participants notice this change, they will increase their attention. If children do not react differently to the test stimulus, they have probably not noticed the change from the habituation stimulus. Habituation paradigms are widely used in infant research, because they do not require a verbal response from the participant.

In the switch paradigm, infants are habituated to two label-object combinations. On each trial, the infant sees one object on the screen and simultaneously hears an auditory label. Both label-object combinations are presented in this way on an equal number of trials. When the habituation criterion is reached, the participant enters the test phase, which consists of one *same* trial and one *switch* trial. On the same trial, one of the label-object combinations from the habituation is repeated. On the switch trial, the infant sees the object from one combination, paired with the label from the other combination. If infants have learned the association between the labels and the objects, they will display a switch effect: looking longer during the switch than the same trial. If they have not learned the association, there will be no switch effect, because both test trials contain a familiar object and label. Figure 1 gives an example of the phases in this two-word version of the switch task. In the simpler one-word version of the switch task, children learn only one label-object combination during habituation. On the switch trial children see the object from the habituation phase, but hear a (slightly) different sound.

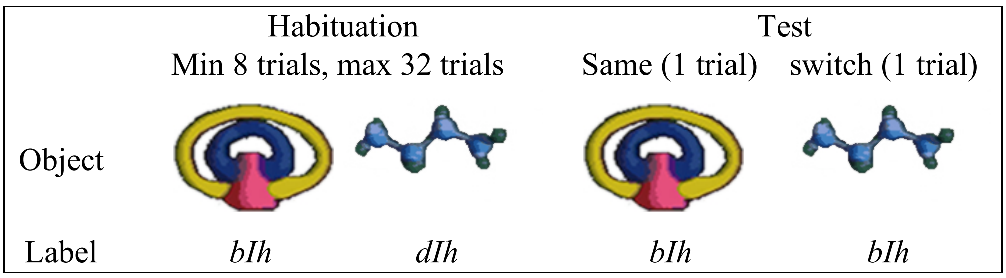

Figure 1 The phases in the two-word version of the switch task. The objects presented here are the 'crown' and the 'molecule' object as used by ((Werker, Fennell, Corcoran, & Stager, 2002), pictures obtained from www.infancyarchives.com).

In their first demonstration of the switch paradigm, Werker and colleagues showed that 14-month olds are able to simultaneously learn two label-object combinations if the words are *lIf* and *nim*. In contrast, children of this age do not show a switch effect when they are habituated to the minimal pair words *bIh* and *dIh* (Stager & Werker, 1997), nor in the one-word version of the task (e.g. when they are habituated to an object paired with *bIh* and on the switch trial heard the label *dIh* for the same object). The switch effect does appear for these words when the visual stimulus is a checkerboard, indicating that the children's apparent inability to notice the switch is not due to speech perception difficulties. Infants' inability to notice the switch has been replicated in English for the phonotactically well-formed words *bIn* and *dIn* (Pater, Stager, & Werker, 2004; Werker et al., 2002) as well as in Dutch (Fikkert, In press), and for different stop consonant contrasts ((Pater et al., 2004): *bIn-pIn* and *dIn-pIn*). Interestingly, 8-month old children do react to a switch between *bIh* and *dIh* in the one-word version of the switch task, which shows that these infants do not yet link the sound to the object, but simply depend on their speech discrimination skills (Stager & Werker, 1997). Older children of 17 and 20 months of age do notice the switch as well, as do14-month olds with larger vocabularies (Werker et al., 2002). Infants' inability to display a switch effect for the words *bIn* and *dIn* is thus a brief developmental stage at the onset of their vocabulary spurt.

<u>Infants' failure in the switch task: due to problems in word learning or recognition?</u>
Because children do not notice the switch between *bIn* and *dIn* and other minimally different words, one could argue that infants' representations of these words are incomplete. The representations could lack, for instance, a specification for the place of articulation of stop consonants, such that the words *bIn* and *dIn* have the same representation. In that case, hearing the word *bIn* does not violate the child's expectation when she sees the 'dIn' on the screen.

There are several reasons why these representations could be incomplete and each of these can be related back to the three stages in the word learning process defined in Table 1: Attending to, encoding of and storing of information. (Werker et al., 2002) state that in a word learning task, which is presumably computationally demanding, infants "cannot attend as closely to the fine phonetic detail that is available, leading to incomplete information uptake." (p. 24) If children do not attend to all information, it cannot be encoded and stored either, leading to incomplete representations. Werker et al. thus seem to locate children's problem in step 1 of the word learning process.

It is also possible to locate the children's problem in the encoding step, step 2. According to the Linguistic Perception model (Boersma, Escudero, & Hayes, 2003; Escudero, 2005) children's language perception system initially encodes only one feature at a time and is unable to integrate several phonetic dimensions. This leads to stored representations with single cues, rather than integrated phonemes. Cues that are louder and produced more consistently in the language environment are most likely the cues that infants encode and store (Curtin et al., In press). If two words differ in a cue that is not easily encoded, children will not represent the difference between the words and thus not notice the switch.

The third option is that children do not experience problems in either attending to the signal or encoding the information, but in storing information. In that line, (Swingley, 2003) proposes that infants attend to the relevant aspects of the speech signal, but lack the computational abilities to store all this information. (Fikkert & Levelt, 2008), see also (Fikkert, In press), assume that infants encode all information, but are restricted by phonological principles as to what they can store in word form representations. Their account focuses on place of articulation features and they propose that children only store the features of the vowel and not the consonants. Also, listeners only specify the features 'labial' and 'dorsal', but not 'coronal'. As a consequence, the words *bIn* and *dIn*, with the same coronal vowel, share the representation ∅ (i.e. underspecified). Because every perceived place of articulation feature is in line with this underspecified representation[2], perceiving the word *bIn* is all right when the object 'dIn' appears on the screen. This means that children will not notice the switches between these words.

Within these explanations that share the view that infants' lexical representations are incomplete, there is a natural division between the computational resources accounts of Swingley and Werker et al. on the one hand, and the principled accounts of Escudero and Fikkert and Levelt on the other. All these explanations are confirmed by the absence of a switch effect. In what follows I will refer to the general position that the absence of a switch effect is due to incomplete representations, whatever has caused them to be incomplete, as the *strong* interpretation of the switch design. The lack of a switch effect does not necessarily imply that infants have incomplete representations. In what follows, I will call this the *weak* interpretation of the switch task and two accounts within the *weak* interpretation are discussed here.

(Yoshida, Fennell, Swingley, & Werker, In press) propose that infants' word representations are not incomplete, but still noisy and phonetic, rather than binary and phonological. Children thus have a rough idea about how the word *bIn* should sound, but do not find the word *dIn* a categorically incorrect realization. Although they might notice that *dIn* is a deviant realization of the word *bIn*, the deviation is not large enough for them to regard *dIn* as a different word and thus display a switch effect.

Yet a different approach to explaining infants' failure to notice the switch is assuming that their representations are complete, but that they experience problems in the word recognition process. In the adult speech perception literature, it has been shown that adults experience lexical feedback to the perceptual level. As a consequence of this top-down process, listeners perceive an ambiguous phoneme in accordance with their lexical context such that they rather hear a word than a non-word (e.g. (Ganong, 1980)). A similar process could play a role in young infants' performance in the switch task. On the switch trial, infants see an object they know a label for and they will therefore expect to hear this label (e.g. *bIn*). When they then hear a slightly different label (e.g. *dIn*)*,* they might do what adults do and reinterpret the signal such that they recognize the expected word form *bIn* and do not notice the switch. This explanation of infants' failure to notice the switch is parsimonious, in that it bears on processes that have been shown to play a role in adult speech perception as well.

Infants succeed in the switch task in the presence of contextual support: explanations

In a series of studies following up to this absence of a switch effect, Fennell and collaborators have found that infants do show a switch effect for *bIn* and *dIn* in the one-word version of the switch task if 1) they are familiarized with the object prior to the experiment (Werker & Fennell, In press); 2) the word is presented in a sentence context (Fennell, 2006); or 3) the habituation phase is preceded by three trials with familiar word-object combinations, such as the word *shoe* with a shoe (Fennell,

[2] Or, to be more precise, no place of articulation feature is not in line with the underspecified representation.

Waxman, & Weisleder, 2006). Infants also succeed in the two-word version of the switch task if the words are familiar to the infants prior to participation, *ball* and *doll*[3] (Fennell & Werker, 2003). These results all show that infants are able to attend to small detail in a word learning task, but only if they are supported by some contextual factors. The computational resource explanation seems best in line with the results from this work.

Fennell and colleagues argue that their manipulations take away some of the cognitive load of learning a label-object combination. When infants already know the object (as in (Werker & Fennell, In press)) they do not have to spend computational resources on encoding that part of the combination. Also, 18-month old infants process familiar words quicker if it is presented in a naming phrase than in isolation (Fernald & Hurtado, 2006)[4] and when a novel label is presented in a referential phrase (Fennell, 2006), infants might therefore need less resources to process the information. In either case, some aspect of the word-learning task requires less computational resources, such that these can be dedicated to a better uptake of information about the word form.

I am not convinced that presenting children with familiar word-object combinations would lower computational demands. In the first place, the habituation trials are exactly the same as in the earlier switch experiments, without direct contextual support. An additional explanation is thus needed for the transfer from the first three trials to the habituation. Also, it could even be argued that seeing familiar word-object combinations increases computational demants, because children might put effort into keeping these words active throughout the task. (Fennell et al., 2006) argue that including three trials with familiar word-object combinations before the trials with the novel word clarifies the referential nature of the labels the children hear and in that way helps the children to store detailed information about the form of the novel word presented thereafter. An alternative account that follows this assertion and is in line with the computational resources explanations is that infants increase the resources put into the word-learning task if the context makes it clear that these combinations are important to learn. Making clear to the children that the labels they hear have a referential status is one way of stressing the importance. In the traditional switch design, nothing encourages the children to learn the label-object associations they are presented with. This lack of encouragement might cause the children to hold back in the effort they put into the task. The explanation that children increase the effort they put into the task in the presence of encouragement would be supported if also encouragement of a non-referential nature would enable children to notice the switch, for example if the single words are interspersed with phrases like *Good job!*, *Good working!* and *You're doing great!*

To summarize this section: some form of the computational resources explanation seems necessary to account for infants' success in the switch task with contextual support. This could be extended with the assumption that infants encode novel word forms in a noisy fashion (Yoshida et al., In press), but in a less noisy fashion if the referential status of the labels is made clear. The computational resources explanation can also be combined with my proposal presented above that top-down processes prevent children from noticing the switch, as more computational

[3] This is a minimal pair in western Canadian English. Note that the vowel is different than in the pair *bIn-dIn*.

[4] On the other hand, other work indicates that only 15-month old girls can learn words embedded in a naming phrase, whereas boys of the same age appear to be better when the words are presented in isolation (Trehub & Shenfield, 2007).

resources might make them more resistant to top-down interference on their speech perception. Explanations stating that infants are in principle incapable to encode (Escudero, 2005) or store (Fikkert & Levelt, 2008) certain information are challenged.

A last explanation is required as to why infants cannot notice the switch when they have to learn two words during habituation, even in the presence of contextual support. For this, I turn again to the adult word recognition literature. It is known that adults are subject to lexical competition effects, meaning that listeners are slower to recognize words with many than with few neighbours (Luce & Pisoni, 1998). Also, toddlers have difficulties learning a new word when it is the neighbour of a familiar word, and learning a new neighbour impairs their recognition of the familiar word (Swingley & Aslin, 2007). If children have to learn the words *bIn* and *dIn* at the same time, lexical competition will occur and word learning is inhibited.

Although the questions regarding the reason of infants' success in some, but not all versions of the switch task will not be solved in the current thesis, it is important to keep in mind for subsequent discussions that the absence of a switch effect by no means implies that infants' representation of word forms are necessarily incomplete.

1.1.b Two-alternative forced choice paradigms

The second type of experimental paradigm commonly used in research on infants' knowledge of word forms is a two alternative forced choice (2afc) task with looking direction. In a forced choice task, a participant has to choose one of several answers in response to a stimulus. In this specific implementation used to investigate word recognition, the participant sees two pictures, hears one word and it is analyzed which picture or pictures she looks at.

(Cooper, 1974) was the first to investigate adults' eye movements in reaction to auditory information and found that they look at the most related of 9 pictures upon hearing an auditory stimulus. (Thomas et al., 1981) showed that infants of 13 months old are already able to direct their eye gaze towards one out of four objects upon hearing its familiar word, but that 11-month old children do not yet succeed in this task. One potential problem for the younger children in this study could have been that they did not know the words they were tested on very well (Thomas et al., 1981). (Golinkoff et al., 1987) simplified the task by showing only two instead of four potential referents, such that it became a 2afc design with on each trial one target and one distracter. It is in this form that the method has been used most to investigate infant word recognition. Two measures are used in these investigations: the speed with which infants switch from the distracter to the target after hearing the word for the target, and the time they spend looking at the target, compared to the time they spend looking at the distracter. In chapter 2 I will discuss these and other possible measures in more detail.

One of the first studies using the 2afc method to test infants' word form recognition abilities is (Fernald, Pinto, Swingley, Weinberg, & McRoberts, 1998). They found that children of 24 months old switch their eye gaze from the distracter to the target picture already before the end of the word, indicating that they can recognize words upon partial information. They only do this if the referent is unambiguous before the coda (e.g. *doll* when a 'ball' and 'doll' are on the screen), but not when the referent becomes disambiguated by the coda (*doll* when a dog an doll are on the screen) (Swingley, Pinto, & Fernald, 1999).

In this way, not only infants' knowledge of familiar word forms can be tested, but also the recognition of novel word forms. In that respect, (Schafer & Plunkett,

1998) taught 15-month-old children two dissimilar novel words (*bard* and *sarl*) for two novel objects. On all test trials, the objects were presented side-by-side. The results showed that infants fixated the target object longer upon hearing the name of the target than when they heard the name of the distracter.

In a recent study, (Yoshida et al., In press) showed that 14-month old children that are habituated on the labels *bIn* and *dIn* as for a two-word switch task, can discriminate between these words in a 2afc task: they look longer at the 'bIn' than at the 'dIn' if they hear *bIn*. It must be remarked that they only did this for the first 4 of the 8 test trials and that the difference in babies' looking at the target over the distracter was small compared to other studies, but given infants' consequent failure on this contrast in the switch task, it is important to consider this result in more detail, as well as the differences between the switch task and the 2afc task.

If children's inability to notice the switch were caused by incomplete representations, they would not be able to tell the words *bIn* and *dIn* apart in a 2afc task. The result from Yoshida et al. therefore illustrates that infants' inability to notice the switch is not caused by incomplete representations, but by an aspect of the word recognition process. My explanation of infants' inability to notice the switch, namely that top-down processes cause them to hear the word form they expect to hear, is in line with the result from Yoshida et al. In a 2afc task, children see both the *bIn* and the *dIn* on the screen. Therefore, infants have no prior expectation as to which word they will hear and let speech perception therefore guide their word recognition. Yoshida et al. argue that an important property of the 2afc task is that it involves a direct comparison between the two objects and therefore the two word forms. If children have noisy representations of the words *bIn* and *dIn*, a token of the word *dIn* will not be an incorrect instance of the child's representation of *bIn*, but it will be a better example of the word *dIn*. In either explanation, seeing two objects on the screen enables children to display more sensitivity to phonetic detail than they can show when only one object is displayed.

It is in this respect it is also important to consider that in the switch task, the trial on which word recognition is tested, the switch trial, has exactly the same format as the trials on which the word is learned, the habituation trials. It is thus possible that on the switch trial in the switch task, the child is still in a learning mode and tries to accommodate what she hears into what she has learned previously. In a 2afc task on the other hand, the teaching and test trials have a different format. It is possible that this change in context makes the child aware that the learning process is over and she has to apply her skills now. If this is correct, the switch task gives us more insight in the learning process of novel words, whereas the 2afc task taps into the recognition of the (temporary) end state of this process.

A last remark I want to make is that if the child sees the objects bIn and dIn side by side on the screen and is requested to look at the *bIn*, she will realize that this label refers to only one of the objects. The work of Fennell, discussed previously, has shown that a referential context helps children to notice the switch. Together with this result from Yoshida et al., one could conclude that children do encode and store complete word forms also in the absence of contextual support, but only actively attend to the details in the presence of contextual support.

Either of these four explanations or a combination can be correct in accounting for the difference in results between the switch and the 2afc paradigm, but whichever interpretation is correct, all suggest that the 2afc paradigm might be more sensitive to children's early representations than the switch paradigm. This conclusion will guide the choice of the paradigm for my study. I will come back to these issues in the final discussion of my thesis.

The mispronunciation two-alternative forced choice paradigm

The basic interpretation of the 2afc paradigm that infants look at the object whose word form representation best matches the incoming auditory label, forms the basis of the mispronunciation 2afc (MP-2afc) paradigm, introduced by (Swingley & Aslin, 2000). In a MP-2afc task, children hear correct pronunciations (CP) and mispronunciations (MP) of words. Children display a MP effect if they look less at the target (e.g. a 'baby') upon hearing the MP, (e.g. *vaby*) than the CP, and a MP effect shows that the CP is a better match of infants' word representation than the MP. The phonetic detail in which the MP differs from the CP must then be included in some form in the child's representation. Results from the MP-2afc paradigm often suggest that children do look at the target object, even if the label is mispronounced (among others, (Swingley & Aslin, 2000)) but the data from (Ballem & Plunkett, 2005) show that if children hear a MP for the first time, they do not look longer at the object with the mispronounced label than at the distracter, whereas they fully accept the MP as correct upon later occurrences. This suggests that children learn to connect the mispronunciation to the target after they are at first unable to do this.

As with the switch paradigm, it is important to consider what can be concluded from the absence of a MP effect in the MP-2afc paradigm. The *strong* interpretation is again that children do not represent the detail in which the MP differs from the CP. This strong position can be defended, because the child has to choose between two words and will thus be less subject to the top-down interference that might come with seeing only one object on the screen. On the other hand, some top-down interference could still occur, since infants will expect to hear the word for one of the objects she sees on the screen. In that case, a mispronunciation might be perceptually repaired to match one of the expected word forms. A *weak* interpretation of the absence of a MP-effect is thus possible as well. On the other hand, it can be argued that perceptual repair takes some time, such that the child can be expected to start looking at the denoted object slightly later if the label is mispronounced.

Now, lets turn to the results. Infants of 18 months and older display a MP effect for word initial consonants in familiar words (Bailey & Plunkett, 2002; Swingley, 2003; Swingley & Aslin, 2000), as do 14-month olds (Ballem & Plunkett, 2005; Swingley & Aslin, 2002). Nineteen-month old children also experience a MP effect for word-medial consonants (e.g. CP *baby*, MP *bady*, (Swingley, 2003)). (Mani & Plunkett, 2007) found that 15- 18- and 24-month olds are sensitive to MPs in vowels as well as consonants, with the sensitivity to consonant MPs being larger for the 15-month olds, but not the older children. Also novel words have been tested in this paradigm, with a MP effect for onset stop consonant (Ballem & Plunkett, 2005)[5] as well as gross vowel mispronunciations (Mani & Plunkett, 2008).

[5] The interpretation of these results is rather tentative, as children did not display any Naming effect on the CP trials in the first block, showing that they had not learned the labels at all. In the second block, children reacted correctly to the CPs but not to the MPs of the novel words. The difference in looking between the CPs and MPs was, however, not significant.

1.2 SPECIFIC DETAILS IN INFANTS' WORD REPRESENTATIONS

In the MP-2afc studies described in the previous section, all kinds of mispronunciations were tested without focussing on specific contrasts.In a first attempt to test whether the kind of mispronunciation would affect children's performance, (Swingley & Aslin, 2002) tested half of their 14-month old participants on large MPs (two or more phonological features were changed, e.g. MP *raby* for CP *vaby*) and half of them on small MPs (one phonological feature was changed, e.g. MP *vaby* for CP *baby*). There was a MP effect in both conditions and no difference between the conditions. With the exception of (Pater et al., 2004), all switch studies discussed so far only focussed on the contrast between words with the onset consonants *b* and *d*.

(White, Morgan, & Wier, 2005) tested 19-month olds' sensitivity to the severity of mispronunciations in a more systematic way in an adapted version of the MP-2afc task. They presented the children on each trial with a familiar object for which the children would know a word and an unfamiliar object for which the children did not know a word, for instance a 'car' and a 'barrel'. During the test, five word-types were presented: A CP of the familiar word (e.g. *car*), a MP of the familiar word with a one-, two- or three-feature change of the onset consonant (e.g. *par, dar,* or *var* respectively) or the (unfamiliar) word for the unfamiliar object (e.g. *barrel*). The one-feature change was always a change in place of articulation, the two-feature change was a change in place of articulation and voicing, and the three-feature change was a change in place of articulation, voicing and manner. Children looked at the familiar object when they heard a CP and looked at the familiar object, but less, when they heard a one-feature MP. The infants were indecisive where to look for two-feature MPs and looked at the unfamiliar object when they heard a three-feature MP or the unfamiliar word. The research I will review below has focussed on even more specific kinds of MPs. First, I discuss studies that show that infants' sensitivity can be asymmetric, with a change in one direction being noticed, but not the change in the reversed direction. The second line of research has focussed on F1 and F2 in vowels.

1.2.a Asymmetries in infants' word recognition

(van der Feest, 2007) tested 20- and 24-month old Dutch toddlers on their recognition of MPs in the Place of Articulation of onset stop consonants in a MP-2afc task with familiar words. The children heard MP stimuli with the coronal onset replaced by a labial (e.g. MP *boos* for CP *doos* (box)) or vice versa (e.g. MP *top* for CP *pop* (doll)). Both 20- and 24-month olds had a MP effect when labial stops were mispronounced as a coronal, but not vice versa. The change from a labial to a coronal consonant is thus more important to these children than the reversed change.

(Fikkert, In press) replicated 14-month olds' inability to notice the switch between *bIn* and *dIn* with Dutch infants in the one-word version of the switch design. She also showed that the infants do not notice the switch between *dIn* and *dOn*, but that they do have a switch effect for the words *bOn* and *dOn*[6] in both directions and for the words *bIn* and *bOn* only if they are habituated on the word *bOn*. Fikkert gives an explanation of these results that bears on abstract and underspecified representations (as in the Fully Underspecified Lexicon model by (Lahiri & Reetz,

[6] Note that English infants' success in noticing the switch between the familiar words *ball* and *doll* (Fennell & Werker, 2003), while they are unable to notice the switch between the novel words *bIn* and *dIn* could be due to such detailed differences, as well as the familiarity status of the words.

1999). Apart from relating to such theoretical considerations, the results show that very small details affect infants' performance in the switch task.

In addition to the asymmetries, Van der Feest found that MPs in place of articulation generally lead to larger MP effects than MPs in voicing and that children react to MPs in voicing only when they are 24 months old, but not when they are 20 months of age. Van der Feest argues that this is not surprising because voicing contrasts are acoustically less salient than place of articulation contrasts (Smits, Warner, McQueen, & Cutler, 2003). A one-feature change along one dimension can thus be more severe than a one-feature change along a second dimension and some dimensions become important before others. Similar results for vowels are discussed in the next section.

1.2.b Vowel F1 and F2 in infant word recognition

(Mani et al., 2008) focussed on infants' knowledge of vowel features in familiar words. They showed a MP effect for vowel height (or F1) and backness (or F2) in 18-month old English children, and did not find a difference between the two. The same children had no MP effect for vowel rounding, which mainly affects the third formant (F3). These results show that also in vowels some dimensions are more important for infant word recognition than others.

Also (Curtin et al., In press) investigated infants' use of specific phonetic details in vowel contrasts. They used the two-word version of the switch paradigm and showed that 15-month old children notice the switch for the pair *dit*-*dIt* (rhymes with *beat-bit*), but not for either *dit*-*dut* (the latter rhymes with *boot*) or *dIt*-*dut*. A closer inspection of the phonetic properties of their stimuli revealed that the tokens of the words *dit* and *dIt* did not overlap in F1, whereas those of the word *dut* overlapped in F1 with the tokens of both *dit* and *dIt*. On the other hand, *dit* and *dIt* showed considerable overlap in the F2, whereas the pairs *dit-dut* and *dIt-dut* were completely separate in this dimension (see Table 2 for a schematic overview). Since children only noticed the switch for the word pair that had no overlap in F1, Curtin et al. conclude that F1 is a very important dimension for young children in word learning, in contrast to F2. They argue that F1 can be expected to be more important for infants than F2 as the former is generally louder and more stable in natural productions. Also, young infants are more sensitive to differences in F1 than in F2 (Lacerda, 1992). On a more general note, these results show again that children's failure to notice the switch is not absolute, but depends on stimulus-inherent factors.

	F1	F2	switch effect
dit-dIt	Separate	Overlaps	Yes
dit-dut	Overlaps	Separate	No
dIt-dut	Overlaps	Separate	No

Table 2 The separation and overlap of the stimuli for the words *dit*, *dIt* and *dut* used in (Curtin et al., In press) and the observed presence or absence of a switch effect for each of the three word pairs.

1.3 RESEARCH QUESTIONS

The results from Curtin et al. differ from those in Mani et al., since the former only found an effect for F1 and the latter an effect for both F1 and F2. Three explanations of these differences are possible. Firstly, vowel F2 might not yet matter for 15-month olds, whereas it has become of importance by 18 months of age. A second option is that infants have difficulties attending to F2, as suggested by the results from Lacerda. Infants of all ages will then have difficulties in learning about the F2 in novel words, as were used by Curtin et al., but might know about the F2 once words have become more familiar, as those tested by Mani et al. If this is correct, 18-month olds should not show an effect of F2 for novel words and 15-month olds should show an effect for F2 in familiar words. Another potential consequence of the weakness of F2 is that children have noisier representations for F2 or be more willing to let top-down processes overrule their perception of this cue. This leads to the absence of an effect in the switch paradigm. If this is the case the apparent difference between 15- and 18-month olds will disappear if they are tested in the same paradigm.

In the current study I try to narrow these three explanations down to at least two by testing 15-month olds on their sensitivity to the F1 and F2 of the vowels of novel words in a 2afc paradigm. The outcome of this research will therefore give us very detailed information about the development of lexical representations, but also more insight in the workings of the two main paradigms used to investigate infant word recognition.

Curtin et al. did not observe asymmetries for either F1 or F2 within their data[7] and Mani et al. do not report asymmetries either. This is remarkable, because asymmetries for vowels have been observed for infants in speech perception (cf. (Polka & Bohn, 2003) for a review). More specifically, English children are better at noticing a change from /I/ to /i/ than vice versa (Dejardins & Trainor, 1998; Swoboda, Kass, Morse, & Leavitt, 1978). In studies where perceptual asymmetries in infant speech perception have been found, babies are generally presented with a repetition of the same stimulus, which is sometimes substituted by a deviant. The participants' task is to react when they hear the deviant sound. This is very similar to the infants' task in the one-word version of the switch design: they have to react to a new sound stimulus. It is therefore maybe not surprising that asymmetries have never been reported for the two-word version of the switch paradigm: In the task with two words, neither of the sound stimuli is new to the infants and this might mask asymmetries. The second question I address in my thesis is whether the asymmetries observed for vowel perception are apparent in word recognition as well. A positive answer to this question would suggest a strong continuity from speech perception to word recognition.

---

[7] Suzanne Curtin (p.c.)

2 OVERALL METHODS OF EXPERIMENTS 1 AND 2

In chapters 3 and 4 I will report on two experiments that have been conducted to address these research questions, Experiment 1 (E1) and Experiment 2 (E2). These experiments share their stimuli, apparatus, coding procedure and general methods of statistical analysis. These will therefore be described and discussed in this chapter, which serves as a reference for the subsequent two chapters. E1 consists of 3 versions, E1.1, E1.2 and E1.3. The differences between these versions will be specified in chapter 3.

The design of all experiments was such that they consisted of a training phase and a test phase. All phases of the experiment were pre-recorded and presented to the infants on a screen. In the training phase, children were familiarized with two novel objects and were taught a novel word for both (E1) or one (E2) of them. In the video phase of the training they saw videos of a person playing with each novel object. In the naming phase of the training, infants saw pictures of the novel objects accompanied by an appropriate speech sound file. In the test phase the children saw the objects side-by-side and were requested to look at one of the objects by an auditory stimulus (e.g. *where is the [word]?)* It was analyzed whether they increased their looking at the target object after it was being named.

An important concern in the current study was a careful control over the properties of the stimuli. Both (Curtin et al., In press) and (Mani et al., 2008) report the F1 and F2 values of their stimuli. The tokens Curtin et al. used for each of the three words *dit*, *dIt* and *dut*, varied considerable in both F1 and F2, which makes it more difficult to assess which dimension they pay attention to. Mani et al. on the other hand seem to have less variation, but do not mention the magnitude of the values they report, which makes an assessment of these claims impossible. One methodological objective in this study was thus to know beforehand what kind of information the children heard during the training phase and to make sure that they heard exactly the same during the test phase. To that end, it was decided to select only one token per word and present that to the children in all phases of the experiment.

2.1 PARTICIPANTS

To keep the participant population as similar as possible to the one from Curtin et al., the infants invited to participate in this study were between 14;15 (442 days) and 15;15 (472 days) months old, and had Canadian English as their primary language. They had no known hearing problems and were recruited from the CH.I.L.D. (CHild and Infant Language Development) database at the University of Calgary. The parents of all participants gave their informed consent prior to their infant's participation. Participating infants received an *infant speech scientist* certificate and either a T-shirt or a book in appreciation of their participation.

Participants were excluded from the experiment if 1) their parent interfered by talking or pointing during the experiment or stopped the experiment; 2) the child was judged too fussy by the experimenter; 3) the experimenter made an error in the procedure; 4) there was equipment failure; or 5) the participant did not complete a pre-set number of test trials (see below).

2.2 STIMULI

All sound stimuli were recorded from the same female native speaker of Western Canadian English in a friendly, child directed manner. All recordings were made in a sound treated booth at the University of Calgary. Recordings were made with a Sennheiser microphone, via a Behringer Eurorack UB802 soundcard on a Macintosh Intel Core 2 duo laptop, with a sampling frequency of 44100 Hertz, using the program Audacity.

2.2.a Novel objects and picture stimuli

Two store-bought objects (displayed in Figure 2) were used as the novel objects. The 'roundy' object (left in the Figure) was a Twisty® Teether ball from Munchkin, with a diameter of 10.5 cm. The 'tubey' object (right in the Figure) was 10 by 13 cm. Their pictures were taking using a 3.2 megapixels digital Olympus camera. The pictures of the familiar objects and animals (displayed in Figure 3) came from the picture database of Dan Swingley's Infant Language Lab at the University of Pennsylvania, where they had been used in previous studies. All pictures had the same light grey background and the pictures were edited for the specific purposes of this study in the program Fireworks® version 8.

For the naming phase in E1.1 and E1.2, all objects were presented in the centre of the screen, with a height of approximately 60 cm. The same holds for the familiar objects in the naming phase of E1.3 and E2. The novel objects in the latter two had a height of 25 cm and moved slowly from side to side, starting and ending in the centre of the screen. These animations were created in the program Final Cut Pro version 5.0.4 and had the same background colour as the pictures.

On all test trials, pictures were horizontally aligned on a light grey background with a distance of 60 cm between their centre points. The pictures on the filler trials were approximately 25 cm high, those of the target objects 30 cm.

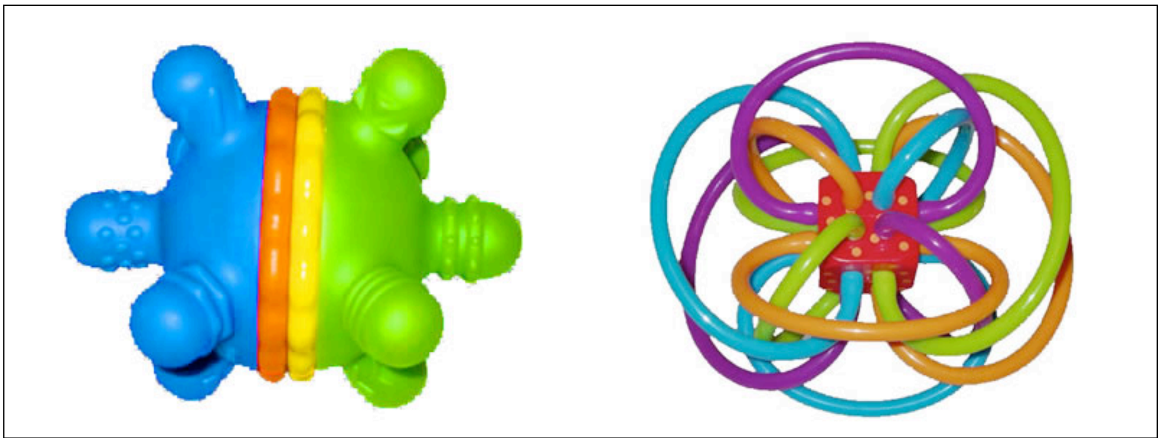

Figure 2 The novel object used in this study. Left: 'roundy', right 'tubey'.

Figure 3 The familiar objects used in this study. From left to right: 'bunny', 'kitty', 'ball', 'bottle', 'car' and 'dog'.

2.2.b Target word tokens

The target words used in this study were the non-words *dit*, *dIt*, *dut* and *dUt* (rhyming with the real English words *feet*, *sit*, *boot* and *foot* respectively). Over 25 instances per word were recorded. In the program Praat (Boersma & Weenink, 2008), I measured the duration of the words as well as their vowels, and the mean F1, F2 and F3 over the second and third quarter of the vowel[8]. One token per word was selected such that the tokens for the pairs *dit-dIt* and *dut-dUt* only differed in F1 and not in F2, and the words in the pairs *dit-dut* and *dIt-dUt* only differed in F2 and not in F1. Informal listening tasks with native speakers of Canadian English revealed that the selected token for *dUt* was perceived as shorter than the other selected tokens (596 ms vs. 657 ms for *dit*; 693 ms for *dIt*; 656 ms for *dut*). Therefore, I lengthened the duration of the vowel in the selected word token *dUt* by reduplicating 7 periods around the midpoint of the vowel /U/, which lengthened the vowel by 19 ms[9]. I also added 40 ms to the pre-burst silence before the coda /t/.

word	duration word (ms)	duration vowel (ms)	F1 vowel	F2 vowel	F3 vowel	
dit	657 ms	366 ms	3.913	13.53	15.946	Bark
			382.553	2215.867	3137.843	Hz
dIt	693 ms	385 ms	5.737	13.955	15.889	Bark
			594.333	2341.894	3111.821	Hz
dut	656 ms	379 ms	3.738	10.545	15.492	Bark
			363.881	1393.797	2936.700	Hz
dUt	655 ms	380 ms	5.721	10.381	15.590	Bark
			592.443	1358.279	2978.628	Hz

Table 3 Duration and formant measurements of the final four tokens used in all experiments and versions, in both the training and the test phases. The formant measurements are the mean over the second and third quarter of the vowel.

Table 3 gives the measurements of the final four tokens as used in all experiments, with a plot of their location in a F1-F2 vowel space in Figure 4. Note that the vowels in the tokens in the pairs *dit-dut* and *dIt-dUt* differed less than 0.3 Bark in F1 (the just noticeable difference for formant frequencies (Kewley-Port, 2001) and the vowels in the pairs *dit-dIt* and *dut-dUt* differed only slightly more than 0.3 Bark in F2. The

[8] This method was chosen in order to avoid coarticulation effects in the measurements, but still have a measure of the overall formant frequencies.

[9] Each reduplicated period was inserted after the original, thereby maintaining the natural pitch and formant contours of the word token.

rounded vowels differed less than 0.3 Bark from each other in F3, as did the unrounded vowels. And although the F3 of the rounded vowels was, as expected, lower than the F3 of the unrounded vowels with the same vowel height, this difference did not exceed 0.454 Bark. These stimuli thus fulfilled the goal that the words in the pairs *dit-dIt* and *dut-dUt* only differed in F1, and the words in the pairs *dit-dut* and *dIt-dUt* mainly differed in F2.

The Euclidian distances[10] between the target word tokens, as computed from their vowels' F1, F2 and F3 in Bark, are given in Table 4. The Euclidian distance between the vowels only differing in F2 (the vowels in the pairs *dit-dut* and *dIt-dUt*) was larger than the Euclidian distance between the vowels only differing in F1 (the vowels in the pairs *dit-dIt* and *dut-dUt*). This is not surprising, given that the auditory distance in F2 was larger than the auditory distance in F1 (see Table 3), and that the tokens that supposedly only differed in F2, differed slightly in F3 as well. As one can see from the comparison between Tables 4 and 5, adding or removing F3 hardly changed these Euclidian distances, though. This very small effect of F3 justifies the focus on F1 and F2 in this research.

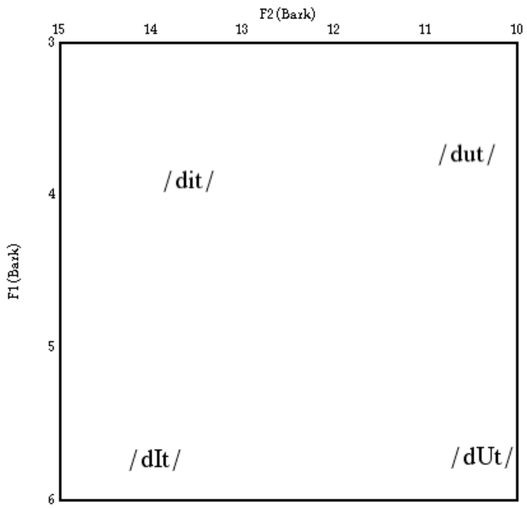

Figure 4 The four tokens for the novel words *dit, dIt, dut and dUt*, in a F1-F2 vowel space.

	dit	*dIt*	*dut*	*dUt*
dit	0			
dIt	1.862	0		
dut	3.077	3.973	0	
dUt	3.695	3.587	1.992	0

Table 4 Euclidian distances between the vowels in the four word tokens, computed over F1, F2 and F3 in Bark.

[10] Computed as $\sqrt{((f1_{word1}-f1_{word2})^2+(f2_{word1}-f2_{word2})^2+(f3_{word1}-f3_{word2})^2)}$

	dit	*dIt*	*dut*	*dUt*
dit	0			
dIt	1.861	0		
dut	3.043	3.953	0	
dUt	3.677	3.574	1.990	0

Table 5 Euclidian distances between the vowels in the four word tokens, computed over F1 and F2 in Bark.

The intensity of these tokens was adapted in Praat until their perceived loudness was judged highly similar. For each of the tokens, a higher version was created with the Manipulation function in Praat. These versions were 9% higher than the originals in E1.1 and 7% higher in E1.2, E1.3 and in E2.

2.2.c Carrier sentences and sound stimuli

For presentation during the naming phase of the training and the test phase of the experiment, the target word tokens, *dit, dIt, dut* and *dUt,* were spliced into carrier sentences. See Table 6 for the carrier sentences in each of the phases and each of the versions of the experiments.

<table>
<tr><th></th><th>E1.1</th><th>E1.2</th><th colspan="2">E1.3</th><th>E2</th></tr>
<tr><td rowspan="2">naming</td><td colspan="2" rowspan="2">-Look at the X! X!
-This one’s a X! X!

(5 s per trial)</td><td colspan="2">filler:</td><td>-This one’s a X! Look at the X! X!

(5 s per trial)</td></tr>
<tr><td colspan="2">novel:</td><td>-This one’s a X! Look at the X! X!
Can you see the X! X! (twice)

(20 s per trial)</td></tr>
<tr><td>test

5 s per trial</td><td>-Can you see the X? X!
-Which one’s the X? X!</td><td colspan="4">-Can you find the X? X!
-Do you see the X? X!
-Where is the X? X!
-Which one’s the X? X!</td></tr>
<tr><td>filler objects</td><td colspan="2">‘kitty’ & ‘bunny’</td><td colspan="3">‘ball’ & ‘bottle’
‘car’ & ‘dog’</td></tr>
</table>

Table 6 The carrier phrases for the naming phase and the test phase used in each version of the experiments, and the used filler objects.

Each carrier phrase was recorded twelve times: three times with each of the four target words. Per carrier phrase, four recording were selected, one with each target word. In the program Praat, the target word was cut off from the carrier at the burst of the /d/ onset and the loudness of all carrier sentences was equated at a level compatible with the four target word tokens. Each target word token was then spliced into each of these carrier sentences.

For all sentences used in the test phase of the experiments, silence was added prior to the carrier sentence, such that the onset of the target word token was at 2.5 s after the onset of the sound file. The higher, single version of the target word token was added 4 s after the onset of the sound file. All files had a total duration of 5 s. The same procedure was used for the sentences in the naming phase of versions E1.1 and E1.2

The sound stimuli for the naming phase of versions E1.3 and E2 consisted of two sentences and one single word token. The silences in between each of the constituent sounds were equally long and the silences before the first and after the last sound were half this length.

The sentences referring to the familiar filler objects were recorded with the correct word and no splicing was applied. The isolated, higher versions of these words were natural recordings as well. The loudness of these sentences was adapted to the same level as the loudness of the carrier sentences for the target words. For the remainder, the stimuli were compiled in the same way as those for the target words.

2.2.d Video Stimuli

In total ten videos were made. In each video, only one object was displayed and maximally one name was given to it. There were eight videos for each combination of the four novel words with the two novel objects, and two videos in which the object did not receive a name. In each video, the upper body and face of a female adult against a beige background were visible and the only colourful entity was the novel object, which the woman held in front of her body, under her face. In the videos in which the object received a name, she said the name of the object six times, embedded in sentence contexts. In the videos without a label, she commented on how interesting the object was in sentences similar to those used in the other videos. During playing, she repeated the same action with the object several times: she twisted one half of the ‘roundy’, or pulled the strings on the side of the ‘tubey’. All videos were strictly scripted, differing only in the object that was used and the name given to the object. Figure 5 shows a still from two videos. See Appendices A.I and A.II for the scripts of the videos with and without names respectively.

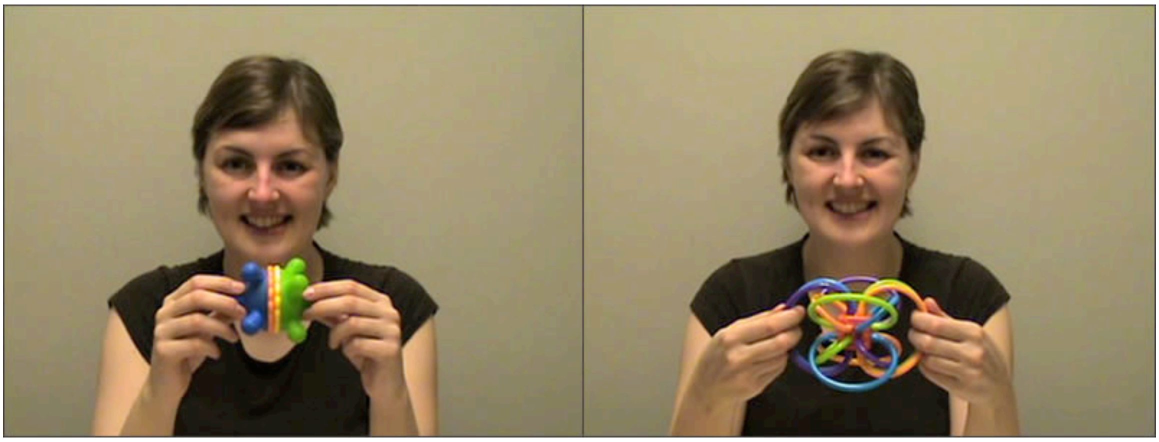

Figure 5 Two stills from the videos. Left: with the ‘roundy’ object. Right: with the ‘tubey’ object.

Videos were recorded in the same room as the target words. The video was recorded on a DCR-DVD92 Sony digital video camera and the sound was recorded independently. For each of the 10 final videos, three takes were recorded and one of these was selected for use in the experiments. The selected target word tokens were spliced into the independent sound recordings of the video with the same word. The pitch of three out of the six occurrences of the target word was manually adapted using the Manipulation function of the program Praat, to ensure that they better fitted the intonation of the carrier sentence. These sound files were added to the original video recordings with the program Final Cut Pro. The independent sound recordings of the videos in which the objects did not receive a name were added to the original video recordings in the same manner, but not manipulated. Native speakers of English judged all videos to look and sound natural.

Each of the videos with a word had a total duration of 41 s. The videos without a word each had a total duration of 38 s. On the projection screen, the objects in the video were 20 cm ('roundy') and 25 cm ('tubey') high.

2.2.e Attention getter

The attention getter was a colourful flashing ball on the centre of the screen against a black background. The attention getter was obtained from Suzanne Curtin's Speech Development Lab at the University of Calgary, where it had been used in previous research.

2.3 APPARATUS

All infants were tested in a sound attenuated small room (177 by 276 cm) with beige painted walls, a white-and-grey board ceiling, a dark grey carpet and a door covered with light green cloth. Two ceiling lamps, one just above the children's head and one to their right, lit the room. These were dimmed to a pre-set criterion for the experiments.

Infants sat either on their parent's lap or on an Expressions™ Evenflo high chair, with their parent on a chair to their left. Children were seated facing the short end of the booth, where the stimuli were presented on a 119.5 by 89 cm SmartBoard projection screen, with their back towards the door. Children's eyes were about 110 cm from the centre of the screen, in an angle of approximately 25°.

A meter behind the position where the children were seated, a white NEC LT245 video projector was hanging from the ceiling. During the experiment, two small green led lights of the projector were constantly on. A round mirror was attached to the wall on the back of the children. This mirror reflected the middle of the screen and was used in subsequent coding to determine the end and start of each trial. Below the screen two curtains in the same beige colour as the walls hid all further equipment from the children's view. A video camera peeked through a small hole in these curtains. As a masking control, the parents wore tight fitting Bose QuietComfort® 2 Acoustic Noise Cancelling® headphones over which Disney music was played.

Stimuli were presented via the Habit 2000 program, version 1.0 (Cohen, Atkinson, & Chaput, 2000), which was run on a Macintosh Power PC G5. The visual and audio stimuli played from digitized files on the computer. The audio stimuli were delivered at 65 dB (+/- 5 dB) via a Crown D-45 amplifier to two Bose 161™ speakers in the testing room, which were attached behind the curtain below the screen. The visual stimuli were projected on the projection screen by the video projector.

The experimenter, blind to the audio stimuli and to the trial type in the test phase, monitored infant's looking via a closed circuit television system from an adjacent testing room. The experimenter initiated every trial by pressing a designated key when the infant fixated the attention getter. Infants were recorded using a DCR-

DVD92 Sony digital video camera. The video recording was used for subsequent frame-by-frame off-line coding.

2.4 PROCEDURE

Each participant was assigned to a condition and a congruent test order prior to coming to the lab.

Prior to the experiment, the procedure and general objective of the experiment were explained to the parents. They signed a consent form and filled out the McArthur Short Form Vocabulary Checklist: Level I of the McArthur-Bates Communicative Development Inventory (Fenson et al., 2000). During this period, the investigator interacted with the child, encouraging her to play with the two novel objects. The experimenter did not name the objects. After the warm-up period, children were brought to a separate testing room and placed either on their parent's lap or in a high chair. Parents received the instruction to not talk or point to their child during the experiment and listened to Disney-music to mask the sound stimuli.

Each trial in both the teaching and the test phase of the experiment started with the attention getter on the centre of the screen. Trials were initiated when the infant fixated attention getter. The experimenter was blind to the target side on the test trials. After the experiment infants received their thank-you gifts.

2.5 CODING

The recordings of the children's face were sampled at 30 ms intervals. I subsequently coded infants' looking direction on each frame with the program Supercoder 1.5 (Hollich, 2005). The recorded image of the mirror served as anchor to indicate the beginning and end of each trial. I was blind to the condition in which the infants had been placed as well as the target side on the test trials.

2.6 ANALYSIS

2.6.a Looking time measures to assess word recognition

The researcher's task in analyzing a 2afc task on looking behaviour is to establish how children direct their eye gaze in reaction to the auditory stimuli. The most common measure to investigate 14- and 15-month olds' behaviour is looking time. It is therefore necessary to compare looking time in the pre-naming phase, before the child knows what the target object is, to looking time in the post-naming phase, after the child has heard the word. An increase in looking in the post-naming phase compared to the pre-naming phase indicates that the child has recognized the word and is called a Naming effect. In a MP-2afc task, the researcher wants to find out whether there is a difference between children's reactions to CPs and MPs. A difference between these two conditions is the MP effect. I will briefly discuss several measures to assess these effects, in order to make an informed choice which to use in my thesis.

Researchers working in the 2afc paradigm are generally interested in children's looking at the target in comparison to their looking at the distracter, thereby ignoring infants' looking at the centre of the screen or entirely away from the screen. Formula 1 gives the simplest way to compute looking at the target in this way.

total looking at target / (total looking at target + total looking at distracter) (1

This is called the Proportion of Target Looking (PTL) measure and yields values varying from 0 to 1. A PTL of 0.5 indicates no longer looking at either object. It has been used in every paper on infant word recognition using the (MP) 2afc method I encountered (cf. Introduction). A second measure takes only the longest look directed at the target and the distracter into account. Looking at the target in the Longest Look (LLK) measure is computed as in Formula 2. A LLK of 0 indicates that the child is not looking longer at either of the objects.

longest look at target – longest look at distracter (2

Although many papers use both measures, a difference in result pattern occurred only for (Schafer & Plunkett, 1998). In that paper, only the LLK measure revealed a significant increase in looking at the target when its newly taught novel name was given. Schafer and Plunkett propose that their post-naming phase of 7 s might have been too long for the children. They therefore suggest that LLK is a better measure of the child's 'stickiness' to objects that match the sound she is hearing and more robust to random behaviour.

The PTL measure will be the primary measure I use in this thesis. The first advantage of the PTL measure is that more of the child's behaviour on a trial than only the longest look is taken into account. A second advantage is that it is a proportional measure and has the same scale (0 to 1) for all children, thereby factoring out some of the inherent variation in absolute looking time between children prior to the analysis. On the other hand, if Schafer and Plunkett are correct in stating that the LLK measure is more robust to random behaviour, it might be more suitable for my experiments. Although the post-naming phase is only 2.5 seconds long, children see the pair of the novel objects at least 8 times during the test phase and this might cause boredom and random behaviour soon after the object has been named. I will therefore analyse the results from the LLK measure as well. In the remainder of this chapter when I only mention the PTL measure, the same applies to the LLK measure.

2.6.b Assessing the Naming effect and the MP effect

When the same word is presented to the children on more than one trial, most researchers use mean scores as the child's overall score for that condition. The advantage of one means score over entering individual scores in the analysis is that it reduces the amount of within-subjects variation, which is always a concern in infant research. I will therefore report such means scores as well. More specifically, the pre-naming PTL will be the average over all trials on which the object occurred. The post-naming PTL will be the average over the trials on which the child was asked to look at the object. A disadvantage of such means scores is that an effect might be partially washed out when the child is repeatedly presented with the same stimulus (e.g(Ballem & Plunkett, 2005). I will therefore analyze the trend within each trial type over the individual trials as well.

As indicated above, a Naming effect is the increase in looking at the target after it has been named, which implies that looking in the post-naming phase is compared to looking in the pre-naming phase. If objects are presented in yoked pairs (e.g. the 'ball' and the 'bottle' are always presented side-by-side), the average overall pre-naming PTL will be 0.5 and the average overall pre-naming LLK will be 0. If the average post-naming measure is higher than this baseline, this indicates that overall the child recognized the words. For assessing children's Naming effect for specific words, it is crucial to correct for object saliency and compare in a repeated measures analysis the post-naming and pre-naming PTL for that object[11].

A MP effect is a difference in looking at the target as a function of the pronunciation of the word form. Because the average pre-naming PTL is only dependent on the object and not on the label, the average pre-naming PTL is equal for the CP and the MP trials. It therefore suffices to compare post-naming PTL on the CP and the MP trials to assess the MP effect. In the case of novel words, this is only feasible if it is established first that there is a Naming effect on the CP trials.

2.6.c Other measures to investigate infants' word recognition

A rather different measure to investigate infants' word recognition in a 2afc paradigm, is testing whether they switch their eye gaze from the distracter to the target object in reaction to the spoken word stimulus and how fast they switch. A MP effect for switch latency is observed when children are slower in switching from the distracter to the target on hearing the MP than upon hearing the CP. A disadvantage of this measure is that only the trials on which the child fixates the distracter at the onset of the stimulus words and then switches to the target can be used in the analysis. Because infants of 14 months old are less likely to initiate a switch (Swingley & Aslin, 2002), this further reduces the data per child and increases variability. For my specific set-ups, this measure is not particularly informative, because it is not clearly defined what object is the target. I therefore do not use this method in my analyses.

In my analyses, I will explore a related measure: the number of looks on each trial at the target and distracter together. In that respect, Schafer and Plunkett did not find a difference in the number of looks depending on whether children heard a correct or an incorrect word. In their experiment, the incorrect word (*geek*) was very different from the correct words (*sarl* and *bard*), such that children might not have been in doubt where to look because they understood that it was not referring to either object. If children, on the other hand, hear a word that is a slight MP of the word, or words, they have just learned, they might be in doubt where to look. This could result in more switches between the objects and thus more looks on the post-naming phase.

The fourth measure I want to explore here is children's total attention in the post-naming phase of CP and MP trials. In the switch paradigm children look longer when they notice an incorrect pronunciation, but in the MP-2afc paradigm look shorter when they notice that a word is pronounced incorrectly. However, infants' absolute looking time is analyzed for the switch paradigm, and often a proportional measure in the MP-2afc paradigm. This leaves open the possibility that children in the MP-2afc paradigm need more time to make sense of the input if they hear a MP, and will therefore look longer at both the target and the distracter on MP trials than on CP trials, although they look proportionally shorter at the target on MP trials.

[11] It is also possible to subtract pre-naming from post-naming looking. A disadvantage of this procedure is that less of the raw data are entered in the analysis.

2.6.d The time window for analysis

An important concern in analyzing looking time data is the time window in which participants can be expected to react to a stimulus. Based on reaction time data showing that 3-month old infants have a minimum latency of 133 to 200 ms in shifting their eye gaze towards a peripheral stimulus (e.g. (Haith, Wentworth, & Canfield, 1993)), with much longer averages, and that older babies are further delayed when they have to disengage from a visual stimulus in order to initiate the shift (Hood & Atkinson, 1993), (Fernald, Swingley, & Pinto, 2001) argued that only looking behaviour from 366 ms after the onset of the stimulus word on should be taken into account in a 2afc task with infants. In most of the subsequent research this criterion is adopted as the beginning of the post-naming phase as well (see (Ballem & Plunkett, 2005) for an exception). However, much research on young infants' visual reaction time indicates that there is considerable variation between children (e.g. (Dougherty & Haith, 1997)). In this thesis I therefore adopt a different procedure to determine for each child individually when she can be expected to react to the stimulus. Prior to every trial, children see a flashing ball in the centre of the screen. This means that at the beginning of every 2afc test trial, children have to switch their eye-gaze from the centre of the screen, which is grey and boring when the attention getter stops, to either the left or the right. The median speed with which a child initiates this switch is her personal reaction time and I take this measure as her personal onset of the post-naming phase[12]. This can be considered a lenient measure, because at the beginning of the trial the flashing ball disappears and children thus do not have to disengage from a present stimulus before looking at the left or the right. Also, children do not have to process the speech signal at the beginning of the trial, something that will presumably slow them down in reacting to the spoken words. Moreover, at the stimulus onset of all test trials, children hear the phoneme /d/, which will not disambiguate the target, such that their reaction to the auditory input of interest can only start slightly later. On the other hand children will look at the target when it is being named on about half of the trials and thus do not need to switch in order to react correctly to the stimulus. (Ballem & Plunkett, 2005) argue that on these trials the reaction to the stimulus might start faster than if a switch in eye gaze is needed. Overall, I think that my measure for the onset of the post-naming phase will be a fair approximation of children's reaction time to the target word. Also, if the measure is either too lenient or too conservative, it is so for all children to the same degree. Considering that (Fernald et al., 2001) found the same result pattern independently of whether they analyzed the results from 100, 200 or 300 ms after the onset of the target word, it is unlikely that the probably differences between children in the onset of the post-naming phase will truly affect the result patterns. That is, however, no reason not to try and improve the method.

Not only the start, but also the end of the time window must taken into account. Fernald et al. consider reactions after 1800 ms after the stimulus onset as random for 18- and 21-month olds. This window was extended to 2000 ms for 14-month-olds by (Swingley & Aslin, 2002) and as such adopted by many researchers afterwards. Here I take 1800 ms after the child's individual starting point of the post-naming phase as the end point of the post-naming phase, thereby allowing each child the same amount of time in the post-naming phase. For the LLK measure, I do not

---

[12] The median is taken to reduce the influence if some very fast or slow reactions that might be due to random factors. (Dougherty & Haith, 1997), for instance, analyze the median, rather than the mean reaction times as well. Trials in which the child did not fixate the center at the beginning of the trial and trials where the child switched after 700 ms were not taken into account.

adopt such a maximum because it is supposedly resistant to long trial durations (Schafer & Plunkett, 1998).

2.6.e Exclusion criteria for trials and participants

Apart from the method of analysis, it should be considered which trials and participants are excluded from the analysis. Regarding the first issue, a trial is necessarily excluded if the participant is not looking at the screen during the post-naming phase. Because infants have to choose which object to look at in the post-naming phase, it can be considered necessary that the children look at both objects during the pre-naming phase. I have not taken over this criterion here, because the target and distracter are the same objects on all test trials with the novel objects and I therefore consider it likely that children will know what the other object is, even if they have not seen it[13]. My exclusion criterion for trials might thus be considered lenient.

Participants need to complete at least three trials per condition (that is, the filler trials, the MP trials and the CP trials) in order to be included in the analysis. This criterion was chosen because the dependent variable is a mean over trials and there is a risk of the mean not being a reliable estimation of children's (variable) performance if it is computed over less than three trials. For E2, this means that children are excluded if they complete less than 9 of the 12 trials and thus have to complete 75% of the trials. For comparison, (Swingley & Aslin, 2002) applied a criterion of 17 out of 24 test trials, which is 60% of the trials, and had no restriction on how many trials per condition could be missed. My exclusion criterion for participants is thus rather stringent.

2.6.f Adapting the alpha level

A last consideration in infant research is the alpha level that is adopted to assess statistical significance. Despite the variability of the participants, researchers stick to the traditional and arbitrary *p*-value of 0.05. Especially with the small amount of participants in the research reported here in particular, this leads to analyses with too little statistical power. I will therefore adopt an alpha of 0.1 for the investigation of the Naming effects and MP effects. Unless stated otherwise, all tests are one-tailed, because there are predictions about the direction of the difference.

[13] This does not hold for the first trial, but I accept all of them nevertheless.

3 EXPERIMENT 1: THE MINIMAL-PAIR MISPRONUNCIATION EXPERIMENT

3.1 INTRODUCTION

In Experiment 1, I implemented the Minimal Pair Mispronunciation paradigm. In this paradigm, children first learn two minimally different words, which differ in two dimensions (e.g. *dit* and *dUt* or *dIt* and *dut*) and are tested on their recognition of these words in a MP-2afc test phase. On the MP trials, children hear words that are minimally different from both words they have just learned (e.g. *dIt* and *dut* if the words *dit* and *dUt* were learned). The crucial aspect of the paradigm is that each MP is minimally different from both newly learned words, but from each word on a different dimension. For instance, *dIt* differs from *dit* in F1, but not in F2, and differs from *dUt* in F2, but not in F1. Investigating infants' looking behaviour on the MP trials and comparing this to the CP trials, can therefore give insight in infants' representation and weighting of these dimensions. There are three possible result patterns.

If (Curtin et al., In press) have revealed 15-month olds' true inability to attend to F2 in novel words, children hearing a MP should look at the object whose name has the same F1 as the MP, and should look as long as when they hear the CP of this word, thereby completely ignoring the difference in F2. For example, they should look as long at the dit when they hear the MP *dut* as when they hear the CP *dit*. Such a result would show that infants' ignorance of F2 in the switch task is not task specific. The second option is that 15-month olds do not react to F2 in the switch task because of task demands. This is confirmed if children have no preference for either object on the MP trials, because the MP mismatches children's knowledge of both words in some respect. The third possibility is that children represent F2 and are able to display this in a 2afc task, but weight F1 heavier. In this case, if children hear a MP, they will look longer at the object whose name matches the F1 of the MP they hear, thereby showing a heavier weighting of F1, but they will look shorter than when they hear the CP, which reveals sensitivity to F2 as well.

Also the effect of perceptual asymmetries for vowels on word recognition can be investigated in this paradigm. Because children are better at noticing a change from a mid-high to a high vowel (e.g. from /I/ to /i/) than vice versa, it is predicted that children show a larger MP effect for F1 if the change goes from a low to a high vowel (e.g. CP is *dIt* and MP *dit* is played). Similarly, it can be investigated whether a change from a front to a back vowel leads to a larger or a smaller MP effect for F2 than a change from back to front.

3.2 METHOD

3.2.a Subjects

The participants in this experiment were 12 children (8 female). An additional 5 infants (all female) were tested but excluded from the analysis due to parental interference (1), experimenter error (1), fussiness (2) and the failure to complete at least 3 of the CP trials (1). The mean age of the participating infants was 460.25 days (approximately 15;3), ranging from 444 to 485 days (14;17-15;27). Parents received results letters after all participants were tested and the study was completed.

Three infants participated in E1.1, four in E1.2 and five participated in E1.3. Eight children were assigned to the *dit-dUt* condition, 4 children learned the words *dIt* and *dut*.

See chapter 2 for information about the recruitment criteria and method.

3.2.b Materials and Apparatus

The training videos that were used in this study were the eight videos with each combination of one of the two novel objects ('roundy' and 'tubey') with one of the four novel words (*dit*, *dIt*, *dut*, and *dUt*). Every participant saw only two of these videos.

In the training and the test phase, a child only heard carrier sentences that had originally contained the words she had to learn. All CPs and MPs were presented in the carrier sentences of these two words. To make this more concrete: the CP *dit* would be presented in carrier sentences that had contained the word *dit*, as well as carrier sentences that had contained *dUt.* The MPs in the test phase, *dIt* and *dut* in this example, were presented in these same carrier sentences. Differences between the carrier sentences could thus not be the cause of differential behaviour on CP and MP trials.

The reader is referred to chapter 2 for more information on the stimuli and all information on the apparatus.

3.2.c Procedure

Each participant was assigned to a word pair (*dit-dUt* or *dIt-dut*) and a congruent test-order. The number of children tested on each word pair and each word-object combination was approximately balanced.

After the general introduction, which is described in chapter 2, the experiment consisted of four parts: the first training phase, the first test phase, the second training phase and the second test phase. The exact properties of these phases are described here for each of the three versions of the experiment.

E1.1

1st Training: First children saw two videos of 41 s long of a person playing with an object and naming it 6 times: one for each of the word-object combinations they were to learn. They then saw two trials of 5 s with a familiar animal on the centre of the screen (a 'kitty' and a 'bunny'), hearing the appropriate name of the animal once in a naming phrase and once in isolation. The training phase was concluded by six trials of 5 s on which infants saw one of the novel objects in the centre of the screen and heard its appropriate name once in a sentence context and once in isolation. Each object was presented on three such trials. Infants heard each novel word 12 times, of which 6 times in the video.

1st Test: For each CP, there were 4 trials in which this word was used (CP trials); for each MP, there were 2 trials (MP trials); and there were 2 filler trials with the 'kitty' and the 'bunny', making a total of 16 trials.

2nd Training: The presentation of the pictures of the novel objects on the screen with the appropriate auditory stimulus was repeated three times for each object. Infants heard each novel word 6 times.

2nd Test: The second test phase had the same properties as the first.

E1.2

1st Training: The first training phase of E1.2 was identical to the training phase of E1.1, apart from the fact that each of the videos was shown twice before the static pictures appeared. Infants heard each novel word 18 times, of which 12 times in the videos.

1st Test: There were 2 CP trials for each word, 2 MP trials for each MP and 2 filler trials with the 'kitty' and the 'bunny', making a total of 10 trials.

2nd Training: The children saw both videos once, followed by six trials with the pictures of the novel objects with the appropriate auditory stimulus. There were three such trials per object. Infants heard each novel word 12 times, of which 6 times in the video.

2nd Test: The second test phase had the same properties as the first.

E1.3

1st Training: Children first saw one video for each word-object combination. They then saw four trials of 10 s long with one of the novel objects moving slowly from side to side and its name being played six times in a sentence context and four times in isolation. Each novel object was presented on two such trials. Each of these trials was preceded by the presentation of a static picture of a familiar object ('ball', 'bottle', 'car' or 'dog'), which was named twice. The training phase ended with a repetition of the two videos. Each novel word was presented to the children 32 times, of which 12 times in the video.

1st Test: There were 2 CP trials for each word, 2 MP trials for each MP and 4 filler trials with the familiar objects also presented to the children in the training phase of the experiment, making a total of 12 trials.

2nd Training: Infants first saw the animations of the objects moving from side to side with the auditory labels, one for each object. Then they were presented with the two videos. Infants heard each novel word 16 times, of which 6 times in the video.

2nd Test: The second test phase had the same properties as the first.

The first trial of the first test phase was always a filler trial, as well as the second trial in E1.3. The next two trials were CP trials, one for each of the novel words. Several test orders were created, in which the pictures were balanced for presentation side, target side on the CP trials and the filler trials and presentation side on the MP trials with each MP word. Children heard each CP and MP equally often in a carrier phrase that had originally contained the one CP as in carrier phrases that had originally contained the other CP. The target side of the words that the carrier phrases had originally contained was counterbalanced as well.

The experiment proper lasted approximately 10 minutes in total.

3.2.d Coding & Analysis

Inspection of the video recordings indicated that children lost their attention and became fussy during the second test phase. It was therefore decided to only analyze the first test phase.

Because in E1.1 and E1.2 infants only participated in 2 filler trials, the exclusion criterion for children was lowered to at least two filler trials instead of three. The exclusion criterion of three completed CP trials was maintained.

See chapter 2 for further details on the coding and analysis.

3.3 RESULTS

3.3.a Individual measure of the start of the post-naming phase.

Because I applied a different criterion than earlier studies to determine the start of the post-naming phase, I first wanted to determine how my new measure relates to the old one. The average start of the children in this study was after 11.75 frames (sd: 1.588), which is 392 ms (sd: 52) and slightly but not significantly later than if (Fernald et al., 2001)'s standard criterion of 366 ms were applied ($t(11)=1.636$, $p_{(2\text{-tailed})}=0.130$). The children's start points ranged from 9 to 15 frames (300-500 ms), which is a range of 200 ms.

3.3.b Recognition of familiar filler words

It was first investigated whether children were able to perform the 2afc task with the familiar filler words. As these were presented in yoked pairs, a PTL over 0.5 or a LLK over 0 in the post-naming phase can be taken as an indication that the children do in general recognize the familiar filler words and can perform the 2afc task. The average PTL in the post-naming phase was 0.569, which is slightly but significantly higher than 0.5 ($t(11)=1.375$, $p=0.098$, higher for 7 of the 12 children). The average LLK in the post-naming phase was 0.12 s, which is not significant above 0 ($t(11)=0.642$, $p=0.267$, higher for 7 children).

The Naming effect for the fillers was thus small and only significant in the PTL measure. It was therefore tested whether the Naming effect for the fillers would be larger when the first trial was not taken into account. Recall that the first filler trial was the first trial of the whole experiment and children might not yet have understood the task. In support of this, we found that the average PTL in the post-naming phase of the first filler trial was 0.43, which is not significantly below 0.5 ($t(11)=0.710$, $p_{(\text{two-tailed})}=0.493$, lower for 7 children), and that the average LLK was -0.497, which is not significantly below 0 ($t(11)=1.458$, $p_{(\text{two-tailed})}=0.173$, lower for 4 children). This indicates that children did not look longer at the correct object on the first filler trial. If the first trial was excluded from the average measure for the filler trials, the average PTL was 0.669, which is significantly higher than 0.5 ($t(11)=2.050$ $p=0.033$) and the average LLK is 0.5778 s, which is significantly above 0 ($t(11)=1.775$, $p=0.052$). These results suggest that children did react correctly on the filler trials that were not the first trial of the test. Furthermore, the difference between the first filler trial and the others was significant in both measures(PTL: $t(11)=1.654$, $p=0.063$; LLK $t(11)=2.033$, $p=0.067$). This shows that the children indeed did not yet know what to do on the first trial of a 2afc task, but understood the task afterwards.

3.3.c Recognition of novel words

To investigate whether the participating infants had learned the two novel words, it was first tested whether the PTL was over 0.5 and the LLK was higher than 0 in the post-naming phase of the CP trials. The average PTL was 0.486, which is lower than 0.5, although not significantly so ($t(11)=0.181$, $p=0.430$, higher for 7 children). The average LLK was 0.221, which is not significantly larger than 0 ($t(11)=0.650$, $p=0.265$, higher for 8 children). This suggests that the infants had not learned the two words. It is, however, possible that children had learned the words, but were confused by the MPs and could therefore not display the correct behaviour anymore after the MP trials. Therefore, we separately investigated children's performance on the first two CP trials, which came prior to the MP trials. The average PTL was 0.536 and the average LLK was 0.111. Neither was significantly higher than expected from chance

(PTL: *t*(10)=, *p*=, higher for 6 children; LLK: *t*(11)=0.486, higher for 7 children). Comparing infants' pre-naming bias to post-naming looking of the two words separately showed that only 3 (PTL measure) or 5 (LLK measure) children had a positive Naming effect for both words they learned. This strengthens the conclusion that the children did not learn two words.

It is a possibility that the children learned only one word, for instance only the word for one of the objects (e.g. the 'roundy' but not the 'tubey'). However, neither the PTL nor the LLK increased significantly in the post-naming phase compared to the pre-naming bias for either object (all $p>0.1$). It is also possible that children only learned the word with the front, unrounded vowel (e.g. *dit* or *dIt*) or the back, rounded vowel (e.g. *dUt* or *dut*). For neither the names with the unrounded vowels, nor the names with the rounded vowels there was an increase in looking at the target object in the post-naming phase (all $p>0.1$). It is however interesting to note that there was an increase in looking at the target for the words with the front, unrounded vowels for 6 (PTL) or 7 (LLK) children, whereas there was an increase for the words with the back, rounded vowels for 8 (PTL) or 9 (LLK) children. There is thus no indication that children systematically learned the word for one object, or for one type of word, although the data suggest that the words with the back, rounded vowels were learned slightly better.

In a last effort to investigate whether or not children had learned at least one word, we selected for each child the 'best' novel word: the word with the largest positive difference between their pre-naming bias and post-naming looking. The other word was called the 'worst' word. Table 7 gives infants' pre-naming bias for the object denoted by their 'best' word, as well as the post-naming looking on the four types of test trials (in both the PTL and LLK measure). We then tested whether children's increase in looking at the object denoted by the 'best' word was significant if they heard it, which was the case indeed (PTL: $t(11)=2.329$, $p=0.020$; LLK: $t(11)=1.870$, $p=0.044$). Interestingly, as can be seen from Table 7, children also increased their looking at the 'best' object if they heard the CP of their 'worst' word or any of the two MPs.

	PTL measure	LLK measure
pre-naming bias	0.461	-0.063
post-naming looking, if 'best' word is said	0.6072	0.449
post-naming looking, if 'worst' word is said.	0.645	0.188
post-naming looking if a MP in F1 of the 'best' word is said.	0.5024	0.2097
post-naming looking if a MP in F2 of the 'best' word is said.	0.5689	0.101

Table 7 Infants' pre-naming bias to their 'best' object, the object denoted by the word for which they displayed the largest increase from pre-naming bias to post-naming looking, as well as the post-naming looking on the four trial types, in both the PTL and the LLK measure.

3.4 DISCUSSION

In this discussion section I will first go into the conclusions that can be drawn from these results and the implications for our general understanding of children's word learning abilities. I will close off with a discussion of the method and analysis, in order to improve this for the second experiment.

The data strongly suggest that the infants did, on average, not learn the two minimally different novel words they were presented with. This is in itself not a new finding (Stager & Werker, 1997; Swingley & Aslin, 2007). Upon closer inspection of the data, it appeared that the children were able to link at least one word to the appropriate object. Interestingly, they also looked longer at this object when they heard the other, supposedly unlearned, word or any of the MPs. We know from other MP-2afc studies that 14-month old infants look longer at the target object than at the distracter if they hear a MP (e.g. *vaby* instead of *baby*, Swingley & Aslin, 2002). Because the CPs and MPs were all monosyllabic with a /d/ onset and /t/ coda, infants could easily have assumed that all words that sounded similar to the one word they had learned referred to this same object. If this interpretation is correct, their representation for the learned word was rather unstable, because the 'worst' word and one of the two MPs deviated from their 'best' word in F1, a dimension (Curtin et al., In press)have shown to be important for infant word recognition. The infants might have been confused about how the word they were learning sounded, because they heard two similar words during training. A more stringent interpretation of these results is that children did not really learn the word for their 'best' object, but only learned that this object was denoted by *some* novel word, without encoding any of the word's specific properties. Because the task only involved test trials with /d/-Vowel-/t/ words, this interpretation cannot be excluded.

What remains to be explained, is why the 15-month old children in the present study cannot distinguish *dit* from *dUt* or *dIt* from *dut* in the 2afc paradigm, whereas children of the same age are able to notice the switch between *dit* and *dIt*, words that only differ in F1 (Curtin et al., In press). Children's failure in our experiment is also unexpected because in the training phase the words were embedded in carrier phrases and the training contained similar trials with familiar objects, two factors that should enhance word learning (Fennell, 2006; Fennell et al., 2006). Also, it should be easier rather than more difficult for infants to keep a minimal pair apart in a 2afc task (Yoshida et al., In press). A first factor to consider is that the participants in Curtin et al's switch task as well as in Yoshida's 2afc task on a minimal pair heard each word at least 40 times before they enter the test phase, whereas the participants in the current study heard each word maximally 32 times. Also, during these habituation phases, children can focus on nothing but the words and the objects whereas our training phase included a person performing actions with an object. Hearing each word less often in a more distracting context can have caused the infants' poor word recognition.

3.4.1 Methodological implications

The first note to make is that (Fernald et al., 2001)'s decision to only analyze children's looking behaviour after 366 ms after the stimulus onset seems fairly accurate. The measure of reaction time I obtained by looking at the speed with which the infants switched from the centre of the screen to the side at the beginning of the test trials after the attention getter had disappeard was very similar to Fernald et al.'s measure, although the majority of the children were slightly slower. It is slightly concerning that the difference between the slowest and fastest child was 200 ms, which is over 10% of the post-naming phase of 1800 ms. If I am correct in assuming that infants' reaction times at the beginning of the 2afc test trials are indicative of the time they need to react to spoken words, the slowest children thus have 10% less time to display their word recognition abilities than the fastest children if a standard criterion is applied. On the other hand, the standard deviation was only 2.8% of the duration of the post-naming phase. Because the measures reported here concern only 15-month old infants, it remains to be seen whether the standard criterion of 366 ms is in line with the observed reaction times of older children as well. An advantage of applying a personal criterion for each child is that age differences, if present, will be automatically corrected for.

A lesson to draw from these results is that the first trial of an experiment does not always give a reliable estimate of 15-month olds' word recognition abilities. (Yoshida et al., In press) have reported similar results for 14-month old infants. In the analysis of experiment 2, I will therefore disregard the first trial prior to the analysis. In earlier research with the same age group (Swingley & Aslin, 2002), including the first trial in the analysis did not seem to affect the overall results, possibly because the number of trials was much larger than in the current experiment and one weak score is then outweighed by many others.

The most important conclusion to draw from these results is that I should not try to teach 15-month olds two minimally different words if the actual objective of the experiment goes beyond that. The interesting properties of the Minimal Pair Mispronunciation paradigm only work if children are able to correctly recognize both objects on the CP trials of the 2afc task. This was not the case here and investigating infants' weighting of F1 and F2 in word recognition proved impossible. It was therefore decided to teach the children only one word in Experiment 2.

4 EXPERIMENT 2: THE MUTIAL EXCLUSIVITY EXPERIMENT

4.1 INTRODUCTION

Because it appeared from Experiment 1 that 15-month olds were unable to learn two minimally different words from the training phase we had set up, but seemed able to connect one name to the correct object, it was decided to only teach them one word for one of the novel objects. This object will be referred to in what follows as the object-with-name. The other novel object was presented as well, but without a name. I will refer to this object as the object-without-name. In the 2afc test phase, the infants see the two objects side by side and hear either a CP or a MP of the name they have just learned. Children will either learn the word *dut* or *dIt*, and either hear a MP in F1, the MP-F1 group, or a MP in F2, the MP-F2 group. As in Experiment 1, three result patterns are possible regarding infant's reliance on and weighting of F1 and F2.

Based on infants' ability to notice the switch between two words when they only differ in F1 (Curtin et al., In press), it is expected that the children will look shorter at the object-with-name if they hear the MP in F1 and will possibly even start looking at the object-without-name. If 15-month olds are not representing F2 for novel words, as the results from Curtin et al. suggest, they will look at the object-with-name when they hear the MP in F2 and look no differently than when they hear the CP. The second possibility is that infants do represent F2 for novel words, but that this cannot be picked up in the switch task. In that case children will look less at the object-with-name when they hear the MP in F2 than when they hear the CP. The results from Curtin et al. are partially replicated if the MP effect is larger in the MP-F1 group than in the MP-F2 group.

In this paradigm, we can investigate the effect of perceptual asymmetries on word recognition as well. Because of infants' better ability to notice a change from mid-high to high vowels, it is predicted that children show a larger MP effect for F1 if the word *dIt* is mispronounced as *dit*, than if the word *dut* is mispronounced as *dUt*. Asymmetries for F2 can be explored in a similar manner.

A MP-2afc task in which the children only know the name of one of the two objects has previously been used by (White et al., 2005). They have shown that this task can reveal clear effects of the number of features changed in the mispronunciations. In the current study, all MPs involve a change in one feature, either vowel height or vowel backness, but based on Curtin et al., it is expected that children of 15-months old treat MPs in F1 as more severe than MPs in F2.

Although 19-month olds are able to perform this task, there is a risk of 15-month old infants failing. (Halberda, 2003) presented children in a 2afc task with a familiar and an unfamiliar novel object, a 'car' and a 'phototube', and asked them to look at the *car* or the *dax*. He found that children of 17 months old look longer at the 'phototube' when they hear the word *dax*, that 16-month olds are highly variable in their looking behaviour and that infants of 14 to 15 months old looked longer at the 'car' upon hearing the novel word *dax* than when they heard the word *car.* This indicates that also the younger infants noticed that the word *dax* is the incorrect label for a 'car', but are unable to infer that it therefore refers to the unfamiliar object.

An important difference between the current study and Halberda's is that the object-without-name is not unfamiliar to the children at the start of the 2afc task, at least not less familiar than the object-with-name. The results from (Werker & Fennell, In press) illustrate that infants of 14 months old are better at learning a word for an object if they are familiar with the object before they have to learn the name. When children see a totally new object, as was the case in Halberda's study, they might not be ready to accept a new word as referring to this object. In our experiment, children have had time to get to know the object-without-name and might therefore be able to interpret a MP as its name. Infants' success in this experiment would therefore not only give insight in their reliance on F1 and F2 in word recognition, but also in the prerequisites for mutual exclusivity as a word learning strategy in 15-month olds.

4.2 METHOD

4.2.a Subjects

The participants in this experiment were 11 infants (4 female). An additional 14 infants were tested but not included in the analysis due to parental interference (2), experimenter error (4), or the failure to complete at least 3 trials within each condition (8). The mean age of the participating infants was 450.64 (sd: 9.902) days (14;24), ranging from 439 to 470 days (14;12-15;13). See chapter 2 for recruitment criteria and method.

Eight infants learned the word *dut*, of which 4 heard the MP in F1 (*dUt*) and 4 heard the MP in F2 (*dit*). 3 infants learned the word *dIt*, of which 1 heard the MP in F1 (*dit*) and 2 heard the MP in F2 (*dUt*).

4.2.b Materials and Apparatus

A custom made 'Word & Object Familiarity' checklist was created for this experiment (see Appendix B). On this checklist parents could fill out whether their child was familiar with the six objects presented to their child during this experiment (a 'ball', a 'bottle', a 'car', a 'dog', the 'tubey' and the 'roundy'), how parents would call these objects to their child and whether their child recognized these word as referring to the objects. Parents were also asked to indicate whether their child was familiar with one of the words *dit*, *dIt, dut* or *dUt*.

In total six training videos were used in this experiment. The four videos with the words *dut* and *dIt* in combination with each of the two novel objects, as well as the two training videos in which the 'roundy' and 'tubey' were only presented to the children and not named.

All carrier sentences a child would hear throughout the experiment had originally contained the word she. Because there were four carrier sentences, the child would hear the same four carriers for all CP and all MP trials. Differences between these conditions therefore cannot be due to differences between the carrier sentences.

The order of the videos was counterbalanced across children, such that some children saw the video with the object-with-name first. Children saw the videos in the reversed order at the end of the training phase. Four orders were created for the single objects on the screen during the naming phase. It was ensured that children always saw each novel object during the first two trials, counterbalancing across children the first object and the presence or absence of the name on the first trial. It was also counterbalanced which familiar object children saw first and which familiar object preceded which novel object.

In the 2afc test phase there were 4 CP trials, 4 MP trials, 4 filler trials, 1 trial in which the object-with-name was presented together with a familiar object and the novel object was named, and 1 trial in which the object-without-name was paired with a familiar object and the familiar object was named. Each test phase started with two filler trials and the two trials in which a novel and a familiar object were presented side-by-side. The order of the latter was counterbalanced across children. Eight lists for the order of the 2afc test trials were created, balancing within each list the pictures for presentation side, the target side on the CP and filler trials and the presentation side on the MP trials. It was also ensured that there were never more than 2 CP trials or 2 MP trials in a row. The 'tubey' and the 'roundy' object were never presented on the same side more than 3 trials in a row, and the target side of the filler and the CP trials was never the same on more than 3 trials in a row.

See chapter 2 for more information on materials and apparatus.

4.2.c Procedure

Each participant was assigned to a novel word (*dIt* or *dut*), a MP condition (F1 or F2) and a congruent test order, balancing the number of boys and girls in each condition.

In addition to the general introduction, which is described in chapter 2, parents filled out the 'Word & Object Familiarity' checklist.

The experiment consisted of two parts, a training phase and a test phase. In the training phase, children first saw two videos: One of a person playing with one of the novel objects and naming it six times, a second with the same person playing with the second novel object without giving it a name. They then saw trials with one of the novel objects moving slowly from side to side for 10 s. If it were the object-with-name, they would hear the word six times in a sentence context and four times in isolation. If it were the object-without-name, the child would be verbally encouraged to look at the object. Each novel object was presented on two such trials. Each of these trials was preceded by the 5 s presentation of a static picture of a familiar object ('ball', 'bottle', 'car' or 'dog') on the centre of the screen, which was named twice in a sentence context and once in isolation. The training phase ended with a repetition of the two videos. The novel word was presented to the children 32 times, of which 12 times in a video. Immediately after the last video, the 14 2afc test trials started.

4.2.d Coding & Analysis

Because only three participants learned the word *dIt*, I will collapse the participants in a MP-F1 (n=5) and a MP-F2 (n=6) group. For the same reason, it is impossible at this point to assess asymmetric effects, caused by the direction of the mispronunciation. This issue will therefore not be included in the results section or the discussion.

See chapter 2 for further information on the coding and analysis.

4.3 RESULTS

4.3.a Individual measure of the start of the post-naming phase.

I investigated again the time children needed to switch from the centre of the screen to one of the sides at the beginning of the 2afc test trials, the measure also taken as the start of the post-naming phase. The average start of the children in this study was after 11.682 frames (sd: 1.779), which is 389 ms (sd: 5.93) and slightly but not significantly later than if (Fernald et al., 2001)'s the standard criterion of 366 ms were applied ($t(10)=1.271$, $p=0.232$). This is in accordance with what was found in Experiment 1.

4.3.b Filler trials

It was first investigated whether children were able to perform the 2afc task with the familiar filler words. Based on the results from Experiment 1, the first trial was not included the average measures. The children's mean PTL was 0.529, which is only slightly and not significantly above 0.5 ($t(10)=0.613$, $p=0.276$, higher for 6 of the 11 children). The mean LLK in the post-naming phase of the filler trials was 0.067, which is slightly and not significantly above 0 ($t(110.371)=$, $p=0.359$, higher for 7 children). This suggests that the children had difficulties with the filler trials.

One possibility is that the children had difficulties with a specific filler word. As Table 8 shows, for all target objects except for the 'bottle', the children (on average) slightly increased their looking at the target object upon hearing its name. Also, according to the parental report 5 of the participants did not know the word *bottle*. It was therefore decided to compute the post-naming looking over the non-first filler trials again, excluding the trials on which the 'bottle' was the target.

Excluding the *bottle* trials increased the post-naming PTL for the filler trials to 0.583, which is slightly higher that the measure obtained previously and significantly higher than 0.5 ($t(10)=1.587$, $p=0.072$). The exclusion of the *bottle* trials increased the post-naming LLK for the filler trials to 0.230, which is higher than the measure obtained with the *bottle* trials, but not significantly higher than 0 ($t(1.155)$, $p=0.137$). Within both the PTL and the LLK measure, the number of children with on average a positive Naming effect was now 8. If we split up these results for the two conditions separately, we find a significant Naming effect in the MP-F2 group (PTL: $t(5)=3.582$, $p=0.008$; LLK: $t(5)=3.044$, $p=0.014$; both: longer looking for all 6 children), but not in the MP-F1 group (PTL: $t(4)=0.583$, $p=0.296$, longer looking for 3 children; LLK: $t(5)=-0.571$, $p=0.299$, longer looking for 2 children).

The results from the PTL measure suggest that some of the children's problems with the filler trials were caused by the trials with the word *bottle* and that they were thus able in principle to perform the 2afc task. The results from the LLK measure indicate that the *bottle* trials were not their sole problem. The results also suggest that the children in the MP-F2 group were better at the 2afc task than the children in the MP-F1 group.

Word	Side-by-side with	Measure	Pre-naming bias	Post-naming looking
Ball	'Bottle'	PTL	0.511	0.583
		LLK	-0.048	0.185
Bottle	'Ball'	PTL	0.489	0.383
		LLK	0.076	-0.297
Car	'Dog'	PTL	0.356	0.402
		LLK	-0.485	-0.333
Dog	'Car'	PTL	0.644	0.744
		LLK	0.403	0.797

Table 8 Children's pre-naming bias and post-naming looking for each of the four filler objects separately, in both the PTL and the LLK measure.

4.3.c Trials with one novel and one familiar object

The second test to see whether children could perform the 2afc task was investigating their looking behaviour in the trials with one novel and one familiar object. On one of these trials, a familiar object was paired with the object-without-name and the child was requested to look at the familiar object. On the other trial, a familiar object was presented side-by-side with the object-with-name and the child heard the novel word of the latter. Table 9 gives children's looking behaviour on the pre- and post-naming phase of both these trial types.

On the trials in which the familiar object was named, infants looked substantially and significantly longer at the object-with-name in the post-naming phase than in the pre-naming phase (PTL: $t(10)=2.222$), $p=0.025$, longer looking for 9 children; LLK: $t(10)=2.509$, $p=0.015$, longer looking for 8 children).

On the trials in which the novel object-with-name was named, children did not look significantly longer to the named object in the post-naming phase than in the pre-naming phase (PTL: $t(10)=0.157$, $p=0.439$, longer looking for 7 children; LLK: $t(10)=0.921$, $p=0.189$, longer looking for 5 children). These results suggest that the children had difficulties recognizing the just-learned novel word in the presence of a familiar object.

Trial	Measure	Pre-naming bias	Post-naming looking
Novel named	PTL	0.421	0.655
	LLK	-0.367	0.630
Familiar named	PTL	0.472	0.482
	LLK	-0.082	0.112

Table 9 Children's pre-naming bias and post-naming looking for the two trials on which a familiar and a novel object were presented side-by-side, in both the PTL and the LLK measure.

Group	Measure	Pre-naming bias	Post-naming looking
All children	PTL	0.516	0.569
	LLK	0.131	0.200
MP-F1	PTL	0.550	0.498
	LLK	0.303	-0.016
MP-F2	PTL	0.488	0.628
	LLK	-0.012	0.380

Table 10 Children's pre-naming bias and post-naming looking at the object-with-name on the CP trials, for the whole group as well as the MP-F1 and the MP-F2 group separately in both the PTL and the LLK measure.

4.3.d CP trials: looking time measures

It was first investigated whether children had an initial bias to look at the object-with-name over the object-without-name. The pre-naming PTL was 0.516, indicating a slight and non significant bias for this object ($t(10)=0.518$, $p=0.308$, bias apparent in 7 children). In the LLK measure, there was a more substantial preference for the object with name, the average LLK was 0.131, but this was again not significant ($t(10)=0.989$, p-0.173, bias apparent in 7 children).

We then tested whether children looked longer at the object-with-name on the post-naming phase of the CP trials than during the pre-naming phase. On average, there was a very slight and non-significant increase in looking at the target object (PTL: $t(10)=0.909$, $p=0.192$, increase for 8 children; LLK: $t(10)=0.288$, $p=0.890$, increase for 7 children).

Splitting up these results for the MP-F1 and MP-F2 group show that the children in the MP-F2 group increased their looking at the object-with-name when they heard the CP (PTL: $t(5)=4.448$, $p=0.003$, increase for all 6 children; LLK:

t(5)=2.010, *p*=0.050, increase for 5 children), but not the children in the MP-F1 group (PTL: *t*(4)=0.476, *p*=0.329; LLK: *t*(4)=0.740, *p*=0.250; both: increase for 2 of the 5 children). To investigate this further, we looked at the four CP trials separately. These are displayed visually for the LLK measure in Figure 6. These Figures show a different recognition pattern over the course of the 2afc task for the two MP groups. The participants in the MP-F2 condition increased their looking at the object-with-name on all four CP trials, whereas the participants in the MP-F1 group did this on the first trial, but then seemed to become insecure where to look upon hearing the CP. This suggests that the type of MP the participant hears influences the recognition of the CP.

4.3.e The MP effect: looking time measures
The crucial test of the experiment was to see whether there was a difference in looking behaviour on the CP and the MP trials and whether this effect was modulated by the MP condition. Table 11 gives the looking time measures of the infants' looking at the object-with-name on the MP trials, their pre-naming bias for this object and their post-naming looking on the CP trials as comparison. Overall, children looked slightly and not-significantly shorter at the object-with-name when they heard the MP than when they heard the CP (PTL: *t*(10)=1.351, *p*=0.103; LLK: *t*(10)=0.772, *p*=0.229).

Comparing the two MP groups on the children's looking at the object-with-name on the MP trials shows that the children in the MP-F1 group looked significantly less at the object-with-name when they heard the MP than the children in the MP-F2 group (PTL: *t*(10)=3.149, *p*=0.006; LLK: *t*(10)=3.748, *p*=0.002), indicating that the MP in F1 is a more severe mismatch of the children's knowledge of the word they just learned than the MP in F2. The five children with the shortest looks at the object-with-name on the MP trials all came from the MP-F1 group. To gain more insight in these patterns, the results were analyzed in more detail within each group separately.

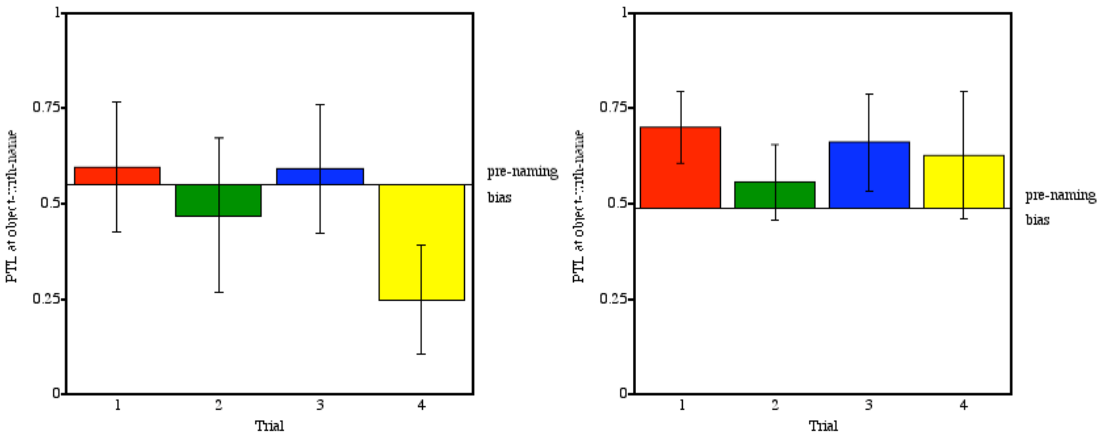

Figure 6 Children's looking at the object-with-name on the four CP trials in the PTL measure. Left: MP-F1 group. Right: MP-F2 group. Error bars display standard error of mean.

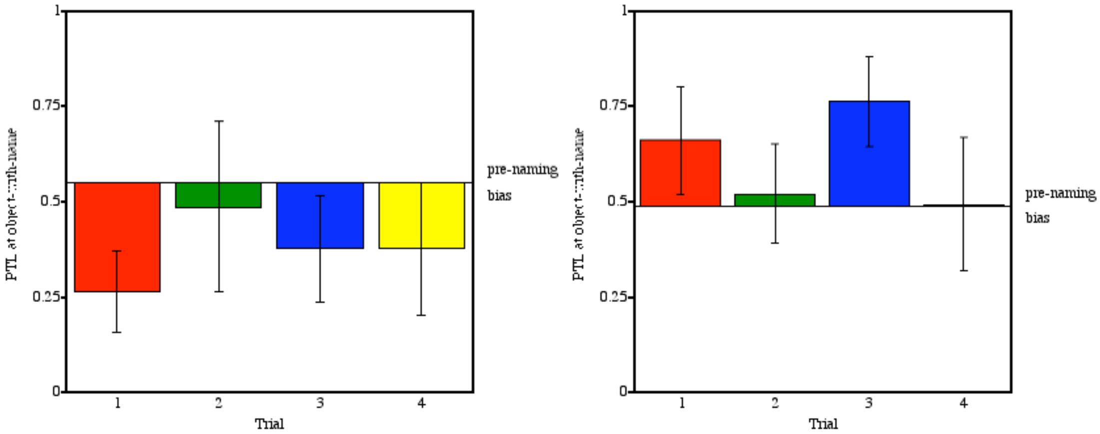

Figure 7 Children's looking at the object-with-name on the four MP trials in the PTL measure. Left: MP-F1 group. Right: MP-F2 group. Error bars display standard error of mean.

Group	Measure	Pre-naming bias	Post-naming looking: CP	Post-naming looking: MP
ALL	PTL	0.516	0.569	0.493
	LLK	0.131	0.200	0.032
F1	PTL	0.550	0.498	0.374
	LLK	0.303	-0.016	-0.493
F2	PTL	0.488	0.628	0.593
	LLK	-0.012	0.380	0.469

Table 11 Children's pre-naming bias and post-naming looking at the object-with-name on the MP trials, for the whole group as well as the MP-F1 and the MP-F2 group separately in both the PTL and the LLK measure.

The children in the MP-F1 group looked substantially, but not significantly shorter at the object-with-name if they heard the MP than if they heard the CP (PTL: $t(4)=1.418$, $p=0.115$; LLK: $t(4)=1.368$, $p=0.122$). Four of the 5 children in this group looked, on average, shorter at the object-with-name when they heard the MP than after the CP. In addition, 4 of the children looked even shorter at the object-with-name after they heard the MP than in the pre-naming phase and for the whole group, this difference between pre- and post-naming looking was significant (PTL: $t(4)=1.991$, $p=0.059$; LLK: $t(4)=2.142$, $p=0.049$).

The children in the MP-F2 group looked slightly shorter at the object-with-name if they heard the MP compared to the CP according to the PTL measure and slightly longer at the object-with-name according to the LLK measure, a pattern followed by 4 children (PTL: $t(5)=0.470$, $p=0.329$, shorter looking for 4 children; LLK: $t(5)=0.350$, $p=0.370$, longer looking for 4 children). Also, most children looked longer at the object-with-name after they heard the MP than in the pre-naming phase (PTL: $t(5)=1.422$, $p=0.107$, longer looking for 5 children; LLK: $t(5)=2.431$, $p=0.030$, longer looking for 6 children).

These results show that the children that heard the MP in F1 had a non-significant MP effect. More importantly, they looked longer at the object-without-name than at the object-with-name when they heard the MP, indicating that they interpreted the MP as the new word for the object-without-name. The children hearing the MP in F2, on the other hand, increased their looking at the object-with-name when they heard the MP.

In a final analysis of the MP trials, we looked at the children's post-naming looking on the MP trials on the four trials separately. These are given in Figure 7. In the MP-F1 group, children looked more to the object-without-name on all MP trials and this effect was largest on the first trial. The children in the MP-F2 group, on the other hand, looked more at the object-with-name on each MP trial, although the effect almost disappeared at the last trial.

4.3.f Test trials: attention and number of switches

As announced in chapter 2, I also investigated children's overall attention and the number of looks in the post-naming phase of the CP and MP trials. Tables 12, 13 and 14 give these numbers for, respectively, all participants, and the MP-F1 and MP-F2 group separately. Because the first test trial with both novel objects side-by-side was always a CP trial, both infants' attention and the number of looks to the first CP trial were largest of all test trials. Comparing the CP and MP trials based on the overall averages in the number of looks and attention would not be very informative, because the averages of the CP trials are boosted by this first trial. Therefore the average over the second and third CP trial and the average over the second and third MP trial were compared. As can be seen from Table 12, children have a slightly larger number of looks on the MP trials than on the CP trials ($t(10)=0.633$, $p=0.270$), and a slightly lower attention ($t(10)=0.036$, $p=0.486$).

Trial type	Measure	Average over trials 2 and 3	1	2	3	4
CP	Looks	2.50	3	2.455	2.545	1.909
	Attention	0.781	0.960	0.792	0.607	0.686
MP	Looks	2.727	2.364	2.455	3	2.455
	Attention	0.779	0.702	0.797	0.760	0.734

Table 12 Children's total number of looks and attention in the post-naming phase of the CP and MP trials. These are the averages for all children.

Trial type	Measure	Average over trials 2 and 3	1	2	3	4
CP	Looks	2.200	2.800	2.000	2.400	1.800
	Attention	0.702	0.945	0.735	0.640	0.695
MP	Looks	2.800	2.400	2.200	3.400	2.400
	Attention	0.816	0.829	0.767	0.865	0.625

Table 13 Children's total number of looks and attention in the post-naming phase of the CP and MP trials. These are the averages for the 5 children in the MP-F1 group.

Trial type	Measure	Average over trials 2 and 3	1	2	3	4
CP	Looks	2.750	3.167	2.833	2.667	2.000
	Attention	0.847	0.973	0.839	0.579	0.679
MP	Looks	2.667	2.333	2.667	2.667	2.500
	Attention	0.747	0.597	0.821	0.673	0.824

Table 14 Children's total number of looks and attention in the post-naming phase of the CP and MP trials. These are the averages for the 6 children in the MP-F2 group.

In a comparison of the groups, we found no difference in the number of looks on the MP trials ($t(10)=0.231$, $p=0.411$), nor in the attention on the MP trials ($t(10)=0.715$, $p=0.246$). From visual inspection of the data, it appeared that the children in the MP-F1 group had more looks and more attention during the MP trials than the CP trials, whereas the children in the MP-F2 group appeared to have less looks and less attention during the MP trials.

4.4 DISCUSSION

The results from Experiment 2 show that children of 15 months old notice a MP in F1 better than a MP in F2. This confirms the findings from (Curtin et al., In press), who found the same difference between these dimensions in the switch paradigm. Because of the small sample size, more data are required to draw firm conclusions about children's sensitivity to F2, as well as about asymmetries. I will first consider the methodological implications of this result and then turn to the relation with previous research on infant word recognition.

4.4.a Methodological implications

An important result that appears from Fennell's work in the switch paradigm (e.g. (Fennell, 2006)), is that children can notice the switch between the words *bIn* and *dIn* in the simple version of the switch task if there is contextual support. It has been argued that contextual support affects the computational resources the child has available for the task, but the current result shows that the exact nature of the detail involved sets the limit as to what contextual support can help children achieve. Apparently, F2 is such a difficult cue, that children do not easily react to it, even in the presence of more contextual support than children in any of the Fennell studies have received.

The crucial methodological question at this point is of course whether the current results should be taken as supporting a strong interpretation of the switch task and the 2afc task. It is at this point possible to argue that children of 15 months old are unable to represent F2, and thus to favour a strong interpretation. One word of caution is that only 11 children have been tested so far. It is thus possible that there is a small effect for F2 in either the looking time or any of the other measures I defined. Given the absence of an effect for F2 in the results by Curtin et al., it is expected that this effect is weak and more statistical power is needed to find it. If an effect for F2 remains absent with the complete sample as well, the strong interpretation of these results will be the easier one. I will discuss here how the absence of an effect for F2 with a larger sample size would relate to earlier accounts within the weak interpretation of the switch and the 2afc paradigm and thus to the position that F2 is represented despite infants' failure to react to it.

One explanation for the absence of an effect for F2 can be formulated within Yoshida et al.'s notion of noisy, phonetic representations of newly learned words. Because of the variability in F2 in English, it could be that infants' representation of F2 is very noisy as well, even though they only hear one token throughout the whole experiment. For the switch task used by Curtin et al., this would mean that children have some knowledge of how F2 should sound for a word, but readily accept deviations as correct as well. Yoshida et al. stressed that knowing a word for both objects on the screen, might help children in performing the task with the words *bIn* and *dIn*: Despite the noisy representations, the incoming auditory information is a better match to one of them. In the 2afc task that children are presented with in this experiment, they only know the name for one object. Because it might be a difficult inference for children to make that a new word refers to an object-without-name (Halberda, 2003), they might readily link the mispronunciation in F2 to the object-with-name, because it is not a completely wrong realization. A second explanation would fit in my top-down account presented in the introduction. If children see one object-with-name and one object-without-name, they might expect to hear the word they know for the object-with-name. Only a deviation in a strong cue could overrule this top-down expectation.

These weak interpretations of the absence of an effect for F2 are still possible if all measures yield a null result, but the position then becomes something in between strong and weak: the information is represented, but infants are yet unable to use it in word recognition.

4.4.b. Relation to the infant word-recognition literature

From the current data, it is clear that a deviation in the F1 of a word 15-month old children know is enough for them to infer that this deviant word must refer to a different object. Such an effect is clearly not present for the F2. In that respect it is interesting to recall that it appears from some studies using the MP-2afc paradigm, that children do not look longer at the target than at the distracter when they hear a MP for the first time (e.g. (Ballem & Plunkett, 2005)). The current data shows that children accept a MP in F2 already the first time they hear it as referring to the object they know a name for. Because the traditional definition of a phoneme is a speech sound that signals meaning distinctions, differences in F1 but not in F2 seem to have a phoneme-like status for English-speaking infants.

From the data it appears that the kind of MP children hear not only affects their behaviour on the MP trials, but their recognition of the CP as well. More specifically, only the children hearing the MP in F2 remained able throughout the whole experiment to recognize the word they had learned as referring to the object-with-name. One possible explanation for this result is that by coincidence, only the six children in the MP-F2 group and not the five children in the MP-F1 group had learned the novel word. This is further supported by the latter group's weaker performance on the filler trials. If the children in the MP-F1 group had not learned the novel word, however, it cannot be explained that they infer that the MP refers to the object-without-name. For example, if the children have not learned that the word *dut* refers to the 'roundy' object, they have no reason to conclude that *dUt* must refer to the 'tubey' object. Yet, the latter is exactly what the children do when they hear a MP in F1. The weak performance of the children in the MP-F1 group can be understood if we recall that children's ability to recognize a familiar word becomes smaller when they are taught a minimally different novel word (Swingley & Aslin, 2007). It could be argued that the same situation occurs in the current experiment when children hear a MP in F1: they interpret this MP as a novel word referring to the object-without-name. Because the MP in F1 is minimally different from the word the children have learned during the training phase, this minimally different ‘new’ word will impair their recognition of the correctly pronounced learned word. Especially because the word learned during training is so new to them, it is likely that it is very sensitive to factors as lexical competition.

In contrast to the results in Curtin et al. and the current study, (Mani et al., 2008) did not find a difference between F1 and F2 when testing 18-month olds’ recognition of the MPs of vowels in familiar words. Because both the current study and Mani et al.’s have been conducted in a 2afc paradigm, the differential findings cannot be ascribed to differences in task demands. So far it appears that older, but not younger children attend to F2, but it remains a possibility as well that infants of both ages are sensitive F1 and F2 for familiar words, but more sensitive to F1 than F2 for novel words. The question about a difference between recently learned and familiar words in the specificity of their representations has been investigated by (Bailey & Plunkett, 2002), who did not find a difference in 18- and 24-month olds’ sensitivity to mispronunciations of early and recently acquired words. This question has not been targeted with specific dimensions, such as F1 and F2, for which there is by now a clear indication that young children treat them differently.

The current findings extend those from (White et al., 2005), who have used the same paradigm as applied here to test 19-month old children, by showing that also 15-month old children can infer that a novel word must refer to the object they do not know a name for. Moreover, the results show that a one-feature MP of the vowel can already cause this effect. White et al. only found this effect when three features of the onset consonant were chanced. Because they only tested one type of one-feature MP, it is as yet too early to draw conclusions about differences between vowels and consonants for lexical access. Investigating this issue is interesting, because it has been suggested that consonants are universally more important to lexical access than vowels (Nespor, Peña, & Mehler, 2003) and that children are therefore more sensitive to consonants than vowels in word learning (Nazzi, 2005). This is clearly at odds with the assumption that children rely more on information they have learned to perceive earlier (Werker & Curtin, 2005). Nazzi tested the hypothesis that consonants are more important in a task where the words were taught in a live interaction. Given that the production of vowels is more variable than of consonants, it is not unlikely that the nature of the teaching situation, rather than an innate preference, has primed the children to focus on the consonants. Although children do normally learn words in natural interactions, a paradigm as used in the current study, with a completely pre-recorded teaching phase and very rigid control over the stimuli, would be a more stringent test of the hypothesis that consonants are inherently more important in infants’ word representations than vowels.

Apart from insight in the ‘form’ side of word learning, the current results also reveal strategies regarding semantic acquisition. When children of 15 months old see a familiar and an unfamiliar object and hear an unfamiliar word, they look longer at the familiar object than when they hear the correct word for the familiar object (Halberda, 2003). The current results suggest that children of this age are able to use mutual exclusivity in word recognition, but only if they are familiar with the object that by exclusion becomes the referent of the novel word. Similarly, (Werker & Fennell, In press) have shown that children are better at learning a novel word for an object if they have been familiarized with the object before learning the word. There thus seems to be a constraint against learning words for unfamiliar objects. Given that 15-month olds know few words and only for very familiar objects, it might be a feasible generalization over their lexicon that familiar objects have names and familiar objects do not. The current data shows that semantic strategies are dependent upon infants' sensitivity to phonetic detail in word forms.

REFERENCES

Bailey, T. M., & Plunkett, K. (2002). Phonological specificity in early words. *Cognitive Development, 17*(2), 1265-1282.

Ballem, K. D., & Plunkett, K. (2005). Phonological specificity in children at 1;2. *Journal of Child Language, 32*, 159-173.

Benedict, H. (1979). Early lexical development: comprehension and production. *Journal of Child Language, 6*(2), 183-200.

Boersma, P., Escudero, P., & Hayes, R. (2003). Learning abstract phonological from auditory phonetic categories: An integrated model for the acquisition of language-specific sound categories. *Proceedings of the 15th International Congress of Phonetic Sciences,*, 1013-1016.

Boersma, P., & Weenink, D. (2008). Praat: doing phonetics by computer (Version 5.0.05).

Cohen, L. B., Atkinson, D. J., & Chaput, H. H. (2000). Habit 2000: A new program for testing infant perception and cognition. (version 2.2.5c). Austin: University of Texas.

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language. A new methodology for the real-time investigation of speech perception, memory and language processing. *Cognitive Psychology, 6*, 84-107.

Curtin, S., Fennell, C. T., & Escudero, P. (In press). Weighting of acoustic cues explains patterns of word-object associative learning. *Developmental Science*.

Dejardins, R. N., & Trainor, L. J. (1998). Fundamental frequency influences vowel discrimination in 6-month-old infants. *Canadian Acoustical Proceedings, 26*, 96-97.

Dougherty, T. M., & Haith, M. M. (1997). Infant expectations and reaction time as predictors of childhood speed of processing and IQ. *Developmental Psychology, 33*(1), 146-155.

Escudero, P. (2005). *Linguistic Perception and Second Language Acquisition - Explaining the attainment of optimal phonological categorization.* Doctoral Dissertation, Utrecht University: LOT Dissertation Series 113.

Fennell, C. T. (2006). Infants of 14 months use phonetic detail in novel words embedded in naming phrases. In D. Bamman, T. Magnitskaia & C. Zaller (Eds.), *Proceedings of the the 30th Annual Boston University Conference on Language Development.* (pp. 178-189). Sommerville Mass: Cacadilla Press.

Fennell, C. T., Waxman, S., & Weisleder, A. (2006). With referential cues infants successfully use phonetic detail in word learning. In H. Caunt-Nulton, S. Kulatilake & I.-h. Woo (Eds.), *Proceedings of the the 31st Annual Boston University Conference on Language Development.* (pp. 206-217). Sommerville, Mass.: Cascadilla Press.

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

Fenson, L., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., & Reznick, J. S. (2000). Short-form versions of the MacArthur Communicative Development Inventories. *Applied Psycholinguistics, 21*(1), 95-115.

Fernald, A., & Hurtado, N. (2006). Names in frames: infants interpret words in sentence frames faster than words in isolation. *Developmental Science, 9*(3), F33-F40.

Fernald, A., Pinto, J. P., Swingley, D., Weinberg, A., & McRoberts, G. W. (1998). Rapid Gains in Speed of Verbal Processing by Infants in the 2nd Year. *Psychological Science, 9*(3), 228-231.

Fernald, A., Swingley, D., & Pinto, J. P. (2001). When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child Development, 72*(4), 1003-1015.

Fikkert, P. (In press). Developing representations and the emergence of phonology: evidence form perception and production. In *Laboratory Phonology 10*.

Fikkert, P., & Levelt, C. C. (2008). How does place fall into place? The lexicon and emergent constraints in the developing phonological grammar. In P. Avery, B. Elan Dresher & K. Rice (Eds.), *Contrasts in phonology: Perception and Acquisition*. Berlin: Mouton.

Ganong, W. F. (1980). Phonetic Categorization in Auditory Word Perception. *Journal of Experimental Psychology-Human Perception and Performance, 6*(1), 110-125.

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology-Learning Memory and Cognition, 22*(5), 1166-1183.

Golinkoff, R. M., Hirshpasek, K., Cauley, K. M., & Gordon, L. (1987). The Eyes Have It - Lexical and Syntactic Comprehension in a New Paradigm. *Journal of Child Language, 14*(1), 23-45.

Haith, M. M., Wentworth, N., & Canfield, R. L. (1993). The formation of expectations in early infancy. *Advances in infancy research, 8*, 251-297.

Halberda, J. (2003). The development of a word-learning strategy. *Cognition, 87*(1), B23-B34.

Hollich, G. (2005). Supercoder: a program for coding preferential looking (Version 1.5). West Lafayette: Purdue University.

Hood, B. M., & Atkinson, J. (1993). Disengaging Visual-Attention in the Infant and Adult. *Infant Behavior & Development, 16*(4), 405-422.
Kewley-Port, D. (2001). Vowel formant discrimination II: Effects of stimulus uncertainty, consonantal context, and training. *Journal of the Acoustical Society of America, 110*(4), 2141-2155.
Lacerda, F. (1992). Does Young Infants Vowel Perception Favor High-Low Contrasts. *International Journal of Psychology, 27*(3-4), 58-59.
Lahiri, A., & Reetz, H. (1999). The FUL speech recognition system. *The Journal of the Acoustical Society of America, 105*(2), 1091.
Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing, 19*(1), 1-36.
Mani, N., Coleman, J., & Plunkett, K. (2008). Phonological specificity of vowel contrasts at 18-months. *Language and Speech, 51*, 3-21.
Mani, N., & Plunkett, K. (2007). Phonological specificity of vowels and consonants in early lexical representations. *Journal of Memory and Language, 57*(2), 252-272.
Mani, N., & Plunkett, K. (2008). Fourteen-month-olds pay attention to vowels in novel words. *Developmental Science, 11*(1), 53-59.
Nazzi, T. (2005). Use of phonetic specificity during the acquisition of new words: differences between consonants and vowels. *Cognition, 98*(1), 13-30.
Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: toward an understanding of the role of rhythm. *Journal of Experimental Psychology. Human Perception and Performance, 24*(3), 756-766.
Nespor, M., Peña, M., & Mehler, J. (2003). On the different roles of vowels and consonants in speech processing and language acquisition. *Lingue e Linguaggio, 2*, 221-247.
Pater, J., Stager, C. L., & Werker, J. F. (2004). The perceptual acquisition of phonological contrasts. *Language, 80*(3), 384-402.
Polka, L., & Bohn, O. S. (2003). Asymmetries in vowel perception. *Speech Communication, 41*(1), 221-231.
Schafer, G., & Plunkett, K. (1998). Rapid word learning by fifteen-month-olds under tightly controlled conditions. *Child Development, 69*(2), 309-320.
Smits, R., Warner, N., McQueen, J. M., & Cutler, A. (2003). Unfolding of phonetic information over time: A database of Dutch diphone perception. *Journal of the Acoustical Society of America, 113*(1), 563-574.
Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature, 388*(6640), 381-382.
Swingley, D. (2003). Phonetic detail in the developing lexicon. *Language and Speech, 46*, 265-294.
Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition, 76*(2), 147-166.
Swingley, D., & Aslin, R. N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science, 13*(5), 480-484.
Swingley, D., & Aslin, R. N. (2007). Lexical competition in young children's word learning. *Cognitive Psychology, 54*(2), 99-132.
Swingley, D., Pinto, J. P., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition, 71*(2), 73-108.
Swoboda, P. J., Kass, J., Morse, P. A., & Leavitt, L. A. (1978). Memory Factors in Vowel Discrimination of Normal and at-Risk Infants. *Child Development, 49*(2), 332-339.
Thomas, D. G., Campos, J. J., Shucard, D. W., Ramsay, D. S., & Shucard, J. (1981). Semantic Comprehension in Infancy - a Signal-Detection Analysis. *Child Development, 52*(3), 798-803.
Tincoff, R., & Jusczyk, P. W. (1999). Some Beginnings of Word Comprehension in 6-Month-Olds. *Psychological Science, 10*(2), 172-175.
Trehub, S. E., & Shenfield, T. (2007). Acquisition of early words from single-word and sentential contexts. *Developmental Science, 10*(2), 190-198.
van der Feest, S. (2007). *Building a Phonological Lexicon: The acquisition of the dutch voicing contrast in perception and production.*
Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word-object associations by 14-month-old infants. *Developmental Psychology, 34*(6), 1289-1309.
Werker, J. F., & Curtin, S. (2005). PRIMIR: A Developmental Framework of Infant Speech Processing. *Language Learning and Development, 1*(2), 197-234.
Werker, J. F., & Fennell, C. T. (In press). Infant speech perception and later language acquisition: Methodological underpinnings. In J. Colombo, P. McCardle & L. Freund (Eds.), *Infant Pathways to Language: Methods, Models and Research Directions*. Hove: Psychology Press.
Werker, J. F., Fennell, C. T., Corcoran, K. M., & Stager, C. L. (2002). Infants' Ability to Learn Phonetically Similar Words: Effects of Age and Vocabulary Size. *Infancy, 3*(1), 1-30.

White, K., Morgan, J. L., & Wier, L. (2005). When is a Dar a Car? Effects of mispronunciation and referential context on sound-meaning Mappings. In A. Brugos, M. R. Clark-Cotton & S. Ha (Eds.), *Proceedings of the 29th Annual Boston University Conference on Language Development.* (pp. 651-662). Sommerville Mass.: Cascadilla Press.

Yoshida, K., Fennell, C. T., Swingley, D., & Werker, J. F. (In press). Fourteen-month-old infants learn similar-sounding words. *Developmental Psychology*.

APPENDIX A.I
SCRIPT FOR THE VIDEOS WITH A WORD GIVEN TO THE OBJECT

OBJECT IN HANDS IN FRONT OF BODY, LOOK IN CAMERA AND SMILE
LOOK AT OBJECT
"OOH, LOOK!"
PICK UP OBJECT WITH TWO HANDS, HOLD IT IN FRONT OF YOU
LOOK IN CAMERA
"THIS ONE'S A [WORD]"
LOOK AT OBJECT
"LOOK!"
LOOK IN CAMERA
"IT'S A [WORD]!"
LOOK AT OBJECT FROM RIGHT AND FROM LEFT
LOOK IN CAMERA
"LET'S PLAY WITH THE [WORD]!"
"LOOK AT WHAT I CAN DO WITH THE [WORD]"
LOOK AT OBJECT
TAKE OBJECT IN RIGHT HAND
PERFORM ACTION
PERFORM ACTION
LOOK IN CAMERA
"CAN YOU DO THE SAME WITH THE [WORD]?"
LOOK AT OBJECT
PERFORM ACTION
PERFORM ACTION
LOOK IN CAMERA AND SMILE
PERFORM ACTION
PERFORM ACTION
TAKE OBJECT IN TWO HANDS, HOLD IT IN FRONT OF YOU
LOOK IN CAMERA
"WHAT A NICE [WORD]!"
LOOK AT OBJECT
LOOK IN CAMERA AND SMILE

APPENDIX A.II
SCRIPT FOR THE VIDEOS WITHOUT A WORD GIVEN TO THE OBJECT

OBJECT IN HANDS IN FRONT OF BODY, LOOK IN CAMERA AND SMILE
LOOK AT OBJECT
“OOH, LOOK!”
PICK UP OBJECT WITH TWO HANDS, HOLD IT IN FRONT OF YOU
LOOK IN CAMERA
“THIS ONE’S NICE”
LOOK AT OBJECT
“LOOK!”
LOOK IN CAMERA
“LOOK AT THIS TOY!”
LOOK AT OBJECT FROM RIGHT AND FROM LEFT
LOOK IN CAMERA
“LET’S PLAY WITH IT!”
“LOOK AT WHAT I CAN DO WITH IT”
LOOK AT OBJECT
TAKE OBJECT IN RIGHT HAND
PERFORM ACTION
PERFORM ACTION
LOOK IN CAMERA
“CAN YOU DO THE SAME?”
LOOK AT OBJECT
PERFORM ACTION
PERFORM ACTION
LOOK IN CAMERA AND SMILE
PERFORM ACTION
PERFORM ACTION
TAKE OBJECT IN TWO HANDS, HOLD IT IN FRONT OF YOU
LOOK IN CAMERA
“WHAT A NICE TOY!”
LOOK AT OBJECT
LOOK IN CAMERA AND SMILE

APPENDIX B
WORD & OBJECT FAMILIARITY CHECKLIST

Word & Object familiarity checklist *F1&F2 in novel words*

Participant code:_______ Date:______ Form filled out by: father / mother /______

	Is your child familiar with such an object/animal?		What do you call this to your child?	Does your child recognize this label?
1		O Yes ⇨	O Ball ⇨ O ____________ ⇨ O The BabySign sign ⇨	O Yes O No
		O No	O I don't label it	
2		O Yes ⇨	O Bottle ⇨ O ____________ ⇨ O The BabySign sign ⇨	O Yes O No
		O No	O I don't label it	
3		O Yes ⇨	O Car ⇨ O ____________ ⇨ O The BabySign sign ⇨	O Yes O No
		O No	O I don't label it	
4		O Yes ⇨	O Dog ⇨ O ____________ ⇨ O The BabySign sign ⇨	O Yes O No
		O No	O I don't label it	
5		O Yes ⇨	O ____________ ⇨ O A BabySign sign ⇨	O Yes O No
		O No	O I don't label it	
6		O Yes ⇨	O ____________ ⇨ O A BabySign sign ⇨	O Yes O No
		O No	O I don't label it	

Does your child know any of the following words?
(e.g. From a second language, or as a name)

O Deet	(rhymes with *feet*)
O Dit	(rhymes with *sit*)
O Doot	(rhymes with *boot*)
O Dut	(rhymes with *foot*)