Early Word Learners’ Ability to Access Phonetic Detail in Well-Known Words*

Christopher T. Fennell
Janet F. Werker
University of British Columbia

Key words: infants, lexical, minimal pairs, phonetic representations

Abstract

Several recent studies from our laboratory have shown that 14-month-old infants have difficulty learning to associate two phonetically similar new words to two different objects when tested in the Switch task. Because the infants can discriminate the same phonetic detail that they fail to use in the associative word-learning situation, we have argued that this word-learning failure results from a processing overload. Here we explore how infants perform in the Switch task with already known minimally different words. The experiment involved the same phonetic difference as used in our earlier word-learning studies. Following habituation to two familiar minimal pair object-label combinations (ball and doll), infants of 14 months looked longer to a violation in the object-label pairing (e.g., label ‘ball’ paired with object doll) than to an appropriate pairing. These results using well known words are consistent with the pattern of data recently obtained by Swingley and Aslin (2002) in which it was found that infants of 14 months look longer to the correct object when the accompanying well known word is spoken correctly rather than mispronounced. We discuss how these results are compatible with the limited resource explanation originally offered by Stager and Werker (1997).

1 Introduction

Phonetic perception becomes language specific by the end of the first year of life (Werker & Tees, 1984; see also Best, McRoberts, LaFleur, & Silver-Isenstadt, 1995; Pegg &...
Werker, 1997; Polka & Werker, 1994; Tsao, Liu, Kuhl, & Tseng, 2000; Werker & Lalonde, 1988). Young infants attend to both native and non-native phonetic distinctions but by the end of the first year of life easily discriminate only those phonetic distinctions that are used in their native language. Having focused on the phonetic variation that carries meaning in their language, it would seem that the infants would proceed to build their lexicon using this knowledge. But, do infants use their refined phonetic discrimination abilities when they begin learning words in the second year of life?

The existing research on children’s ability to attend to phonetic detail in words provides a somewhat confusing picture of whether older children, much less infants, use detailed phonetic information when learning words. On the basis of a series of experiments with minimal pairs (i.e., word pairs that differ by only one phonetic dimension), Shvachkin (1948/1973) provided evidence that, although infants may initially fail to differentiate similar words, they fill out their phonological inventory by the end of the second year of life and no longer confuse minimal pair words. However, Barton (1978, 1980) reported that two-year-olds still confuse minimally different words if they do not already know them well. Eilers and Oller (1976) reported that three year olds confuse minimal pairs in that they treat single feature phonetic substitutions in well-known words as acceptable pronunciation. Similar findings of minimal pair confusion have been reported by Brown and Matthews (1997), Edwards (1974), Garnica (1973), Gerken, Murphy, and Aslin (1995) and Kay-Raining Bird and Chapman (1998), among others. On the other hand, studies with younger infants indicate that infants can attend to fine phonetic detail in word forms. Jusczyk and Aslin (1995) tested 6- and 7.5-month-olds on their ability to notice changes in the initial phoneme of a familiarized word. The older infants listened longer to words to which they were familiarized (e.g., dog, cup), but not to the minimally different foils (e.g., bog, tup). Yet, testing slightly older infants, Hallé and de Boysson-Bardies (1996) discovered that 11- and 12-month-olds do not seem to attend to fine phonetic detail in well-known words. In a previous study, they tested infants on their preference for words that occur with high frequency in speech to children in comparison to low frequency words (Hallé & de Boysson-Bardies, 1994). The infants chose to listen longer to the more common words. Importantly, however, when infants this age were tested with foils that were phonetically similar to commonly occurring words, they also listened longer to the foils over the low frequency words (Hallé & de Boysson-Bardies, 1996). This finding suggests that, in infants of 11–12 months, the representation of known words picked up from the input does not contain fine phonetic detail, that is, that it does not reflect the full range of perceptual sensitivities available to the child at this age.

This inconsistent pattern of results left the question unresolved of whether young word learners have the ability to use fine phonetic differences in word comprehension. Jusczyk and Aslin (1995) consistently stressed that their task requires only listening to word forms without knowledge of word meaning, while Hallé and de Boysson-Bardies (1996) speculated that infants of 11 months failed to notice phonetic detail precisely because they were listening for meaning. This is a difficult claim to evaluate because the tasks used in these studies did not require the infant to understand that the labels referred to objects. To address this ambiguity, a 1997 study by Stager
and Werker used a task that requires the infant to link objects and labels. Because Stager and Werker wanted to test infants as close to the beginning of the word learning period as possible, they chose a task that should be easy for infants. The picture selection and pointing tasks used in the earlier child phonology studies could be too difficult for infants, and arguably even for toddlers. Thus, task difficulty alone may have led to the findings that young children do not attend to phonetic detail. To address this potential confound, Stager and Werker turned to a word-object associative task called the “Switch” procedure that they had developed in collaboration with Leslie Cohen’s laboratory at the University of Texas at Austin.

In the Switch procedure, infants are habituated to two word-object pairings and tested on their ability to detect a switch in the pairing (Werker, Cohen, Lloyd, Casasola, & Stager, 1998). To assess whether infants have learned not only about the words and objects individually, but have linked object A to word A and object B to word B, they are tested using the Switch design, which involves two test trials. On both test trials a familiar object accompanied by a familiar word is presented. On one test trial (the “same” trial) the familiar word and object are presented in a familiar combination; for example, Object A with Word A. On the other test trial (the “switch” trial) a familiar word and object are presented, but in a new combination; for example, Object A will now be paired with Word B. If the infants have learned about the words and the objects but have not learned the associative link, the “Same” and “Switch” trials will be equally familiar, and should attract equal looking times. However, if the infants have learned the link between the specific words and objects, the “Switch” trial, as a violation of that link, should attract greater looking time than the same trial.

Using the Switch procedure, Werker et al. (1998) demonstrated that 14-month-old infants can learn dissimilar sounding labels (e.g., ‘lif’ vs. ‘neem’); however, Stager and Werker (1997) found that infants could not learn phonetically similar labels in this word-object associative task (e.g., ‘bih’ vs. ‘dih’). This was surprising because the earlier work on infants’ refining of phonetic sensitivities during the first year of life would predict that the /b/-/d/ contrast, which is phonemic in English, would be easy for a 14-month-old English-learning infant to discriminate, and presumably to use in word learning. Thus Stager and Werker ran a series of control studies to further investigate why the 14-month-olds failed to apply their language specific phonetic sensitivities to word learning.

In their first set of control studies, Stager and Werker (1997) tested 8- and 14-month-old infants in the single-object variant of the switch task. Infants are only

---

1 This result has since been replicated under many different conditions. Pater, Stager and Werker (under review) used different phonetic contrasts, including a voicing contrast ([b] vs. [p]) and a place + voicing contrast ([d] vs. [p]). Pater et al. also used more valid CVC word forms (‘bin’ vs. ‘din’), rather than the unnatural, at least in English, CV (lax vowel) word forms used in the original Stager and Werker article. In both cases, the 14-month-olds failed to notice the difference in the similar sounding labels. Fourteen-month-olds continue to fail to learn minimal pairs in the switch task even if the objects are made more distinct and longer habituation trials are given (Experiment 2, Werker et al., 2002).
habituated to one word-object combination in this task and the switch trial entails a switch from the habituated label to a novel label. Werker et al. (1998) had shown that infants as young as eight months can succeed in this version of the task (which does not require linking the word to the object, but can be solved as only a discrimination task) when the words are phonetically dissimilar (i.e., ‘lif’ vs. ‘neem’). Stager and Werker tested whether infants of 8- and 14-months could succeed on the /b/-/d/ contrast. The eight-month-old infants in this study looked significantly longer to the switch from ‘bih’ to ‘dih’ (or vice-versa), but once again the infants of 14-months failed. Stager and Werker hypothesized that the eight-month-old infants succeeded in the switch procedure because they treat it as a straight discrimination task while the 14-month-old infants, due to their status as word learners, shift into a word-learning mode when conditions make word-learning possible (even if word learning is not required). They speculated that the presence of a moving, nameable object enables word learning in infants this age and that, once word-learning is evoked, attention to fine phonetic detail is no longer evident. To investigate this hypothesis, they tested 14-month-old infants in the single-object variant of the switch procedure. In this control study, the two similar sounding labels were paired with one visual presentation that was unlikely to be treated as an object by the infants—a stationary, unbounded checkerboard (see Spelke, 1994). Consistent with this prediction, the infants looked longer to the switch in labels under these conditions. Based on all the above findings, it would seem that infants of 14-months only have difficulty accessing phonetic detail when they are placed in a word-learning situation. Why might this occur?

According to one line of reasoning, infants must gradually construct a phonology in which, among other things, the contrastive phonemic inventory of the language is represented. This is believed to be a gradual process which comes about through the establishment of a lexicon (e.g., Charles-Luce & Luce, 1990; Metsala, 1997; Metsala & Walley, 1998). To deal with the apparent inconsistency between the fine phonetic abilities shown in speech perception studies and the mistakes made in early word acquisition, Brown and Matthews (1997) postulate that two separate developmental patterns exist for phonetic versus phonological development. Phonetic development is described as involving “pruning”: The infant begins life with a fairly comprehensive phonetic repertoire and stops discriminating those differences that do not occur systematically in the input. Phonological development requires “building”: The infant must gradually build a phonology on the basis of the phonemic oppositions encountered in building a lexicon. This account assumes then, a representational discontinuity between phonetic perception and phonological acquisition, and could be seen to account for the results reviewed above.

The other class of explanations rejects the notion that the processes of phonetic and phonological development are independent and argues instead for a continuity between phonetic and phonological representations (Swingley, this issue; Werker & Fennell, in press; Werker, Fennell, Corcoran, & Stager, 2002). Here the theoretical arguments are based largely on notions of parsimony. According to this line of reasoning, if the information is discriminable phonetically, it must be available—barring performance limitations—for phonological use.
The explanation put forth by Stager and Werker (1997) posits a specific performance limitation. They suggest that it is the complex nature of word learning that limits use of the available phonetic information. For a novice word learner, forging a link between a label and an object is a computationally demanding task. Thus, the attentional resources available for attending to the fine phonetic detail of the word are limited. This limited resource explanation rests on the assumption that in any difficult task, something has to give (Casasola & Cohen, 2000; Cohen, 1998). In this case, it is the attention to fine phonetic detail that the child sacrifices. It can be argued that this “less is more” approach is adaptive (Newport, 1990). The benefits of reducing information load to more quickly learn the word outweigh the “cost” of inattention to fine phonetic detail. With only a small lexicon, it is likely not necessary to attend to and pick up the fine phonetic detail, as estimates of child vocabularies confirm that children have few phonetically similar words in their early lexicons (Charles-Luce & Luce, 1995).

The resource limitation hypothesis offers an alternative account to the discontinuity theories of why infants might fail to distinguish minimally different words even though they can discriminate those same phonetic differences in a simple speech discrimination task. Specifically, the resource limitation hypothesis assumes a continuity in the underlying representation with the difficulty lying in the access to the phonetic detail. As such it does not require the child to completely rebuild the phonology de novo. It acknowledges that at least the phonetic detail is represented, and suggests that the problem lies in using that phonetic detail in a word learning situation. In this way it maintains the parsimony of a continuity model while offering an alternative account for the data showing that under some conditions infants and children confuse minimally different words.

The resource limitation explanation predicts that infants who have had more experience and more success in word learning should find the act of word learning easier and should thus be able to devote more cognitive resources to accessing the detail in the word form. One proxy that we have used for proficiency in the past is vocabulary size. We have thus predicted that there should be a correlation between vocabulary size in the early stages of word learning and success in the minimal pair word learning task (for a more complete discussion of the relation between vocabulary size and the ability to access phonetic detail, please see Werker et al., 2002; see also Beckman & Edwards, 2000). To examine this hypothesis, Werker et al. (2002) tested infants aged 14-, 17-, and 20-months using the same similar sounding labels and procedure as Stager and Werker (1997).2 As with the previous studies described above, the infants aged 14-months failed to notice a switch in the minimally different labels. The 17- and 20-month infants learned the minimal pair, as indicated by their significantly longer look to the switch trial. Another important finding was the significant correlation in the 14-month group between vocabulary size and success in the task. This indicates that infants aged 14-months with larger vocabularies act

---

2 The procedure and stimuli, while similar, were not completely identical to Stager and Werker (1997). The objects were visually more distinct and the infants had greater exposure to the object-label combinations.
more like older infants and successfully learn similar sounding word-object pairings. These findings support the predictions from our resource-limitation hypothesis.

To summarize thus far, our work indicates that, at the beginning of the word learning period, infants do not use all of the phonetic detail available to them when placed in a word learning situation. They are able to discriminate the difference between the minimally different labels in a speech discrimination task, but cannot use this ability when linking words to objects. It is not until they become more proficient at word learning that they begin to access the detail present in words. We argue that this pattern of results can be explained by our resource limitation hypothesis. The complexity of learning the link between the object and the word form leads to decreased attention to the word form. This hypothesis appears viable based on our own work, but it must be evaluated in relation to other laboratory studies examining similar aged infants’ access to phonetic detail.

Our findings showing that infants 17- and 20-months of age are able to use fine phonetic detail in the “switch” procedure appear at first to contradict the many studies with slightly older children showing confusion of similar sounding words (Barton, 1978, 1980; Brown & Matthews, 1997; Edwards, 1974; Eilers & Oller, 1976; Garnica, 1973; Gerken et al., 1995; Kay-Raining Bird & Chapman, 1998; Shvachkin, 1948/1973). The resource limitation hypothesis can account for this disparate set of findings if one considers the more difficult nature of the tasks used in the studies which report failure. In virtually every case, they have required the child to show a more differentiated understanding of the meaning of a word by pointing to, or otherwise explicitly selecting, the appropriate referent for the word (see Swingley, this issue, for a similar argument). It is possible that under these even more demanding lexical use situations the infant is again taxed, reducing the availability of computational resources necessary to attend to fine phonetic detail (see Werker & Tees, 1999, for a discussion of different levels of difficulty in word recognition tasks).

But how does the resource limitation explanation deal with the studies showing success at minimal pair use at a younger age than shown in the Switch task? In a recent set of studies, Daniel Swingley and Richard Aslin used a visual fixation task to determine young word learners’ ability to attend to correct and incorrect pronunciations of word forms. Swingley and Aslin (2000) presented 18- to 23-month-old infants pairs of objects (e.g., baby and dog) on a computer screen. While viewing both objects, the infant heard either a correct (e.g., baby) or incorrect pronunciation (e.g., vaby) of one of the object labels. The infants’ looking times to the visual “match,” which was the baby in both conditions, were significantly delayed in the mispronunciation condition as compared to the correct pronunciation condition, thus indicating access to the fine phonetic detail in the word forms. The infants also looked longer to the correct picture after hearing the correct pronunciation\(^3\) than after

---

\(^3\) Unlike the Switch task, the visual fixation task is an on-line task with two simultaneously presented choices. The two indicators of success are: shorter latency to look away from the incorrect object and longer looking times overall to the correct match. The habituation phase in the “switch” task leads to the prediction of a novelty preference in the test phase (i.e., longer looking to the incorrect pairing).
hearing the mispronunciation, although the looking time to target in the mispronunciation condition was above chance. These results are consistent with our work with older infants (the 17- and 20-month data from Werker et al., 2002).

However, a new paper by Swingley and Aslin (2002) provides an interesting challenge to our findings with infants of 14 months. Testing infants aged 14-months in the same visual fixation procedure used in their previous work with older infants, Swingley and Aslin included an additional condition. Half of the infants heard a minimally different pronunciation in the mispronunciation condition, as in Swingley and Aslin (2000). The other half of the sample heard a mispronunciation that differed from the accurate label in additional phonetic features. For example, if the target object was “baby” in the mispronunciation condition, half the infants would hear “vaby” as the label (the close mispronunciation condition) and half would hear “raby” (the distant mispronunciation condition). Focusing only on looking time to the match, Swingley and Aslin reported that the infants of 14-months look significantly longer to the match in the correct pronunciation condition than in either the close mispronunciation or distant mispronunciation conditions. Therefore, the 14-month-old infants in this study appear to attend to all the fine phonetic detail in the word form, unlike the same-aged infants in Stager and Werker (1997) and Werker et al. (2002).

Why would the infants aged 14-months in Swingley and Aslin’s (2002) study readily access and utilize phonetic detail while the same-aged infants in our studies fail to access and use similar detail? Two likely explanations should be considered. The first possibility is that the Switch procedure is too difficult for the infants to demonstrate the ability seen in the Swingley and Aslin study. The task demands are indeed much greater in the Switch procedure, as compared to the visual fixation task. In order to notice the violation of the object-label link present in the switch trial, the infant must, at the very least, remember that the word being presented does not “go with” the object. Better yet would be to remember the other object-label combination and compare it to the one being presented during the switch trial. In the visual fixation task, both objects are presented simultaneously, thus reducing the memory load placed on the infant. The infant can continuously check Object B to confirm or disconfirm his hypothesis that Object A goes with the presented label. The reduction in task complexity could free up cognitive resources, allowing the infant to attend to fine phonetic detail. We are currently conducting a study with Daniel Swingley to investigate this possibility. The other possible explanation for the disparate results seen at 14 months concerns the infants’ prior knowledge of the words and objects, and it is this possibility that the current paper addresses.

Swingley and Aslin (2002) used well-known object-label combinations for their stimuli whereas Stager and Werker (1997), as well as Werker et al. (2002), presented the infants with novel object-label combinations. The infants’ a priori knowledge of both the words and objects in the Swingley and Aslin study could account for the difference in results. According to the resource limitation hypothesis, novice word learners have difficulty accessing the detail in newly learned words because of the cognitive complexity involved in mapping a novel label to a novel object. However, this degree of complexity is not present when recognizing a familiar word that has an established
link with its referent. By using known words and objects as the stimuli, the task changes from one of word learning to word recognition, a potentially easier task.

We explored the above effect of a priori word-object knowledge by testing a group of infants aged 14-months using the Switch procedure. The key difference between the current study and past studies conducted in our laboratory is that we presented the infants with two familiar, rather than novel, object-label combinations. As before, the labels differed only in place of articulation. We predicted that the use of familiar objects and labels would reduce the task demands placed on the infant and would, in turn, allow them to attend to and utilize the fine phonetic information present in the minimally different labels.

# Method

## 2.1 Participants

Sixteen 14-month-old infants completed this study, eight girls and eight boys (mean age, 14 months 9 days; range, 13 months 27 days to 15 months 4 days). All subjects were without apparent health problems, were at least 37 weeks gestation, and were exposed to English at least 80% of the time. An additional 10 infants were tested but were not included in the analyses because they were too restless during testing ($n = 9$) or they were not visible to the coder during at least one trial ($n = 1$).

Subjects were recruited through visiting new mothers at the BC Children’s & Women’s Hospital, and through voluntary response to public service announcements. Participating infants were given an “Infant Scientist” t-shirt and diploma.

## 2.2 Stimuli

The audio stimuli were two CVC words that formed a minimal pair: “ball” and “doll”

4 recorded in infant-directed speech (IDS). IDS is effective in gaining and maintaining infant attention (Fernald, 1985; Werker & McLeod, 1989) and in facilitating word learning in infants (Fernald, McRoberts, & Herrera, 1991). The use of IDS also facilitates infant phonetic discrimination (Karzon, 1985). These stimuli differ only in the place of articulation of the initial consonant. An additional, highly dissimilar nonsense label, “neem,”

5 was used during the pre- and post-test trials.

In a soundproof room, a researcher recorded an English-speaking female producing several exemplars of each word in an infant-directed, rise-fall intonational phrase. Final stimuli comprised 10 exemplars of approximately 0.6 s in duration each, with a 1.5 s silent interval between exemplars, resulting in two audio files of 20 s in duration, one for each word.

---

4 These stimuli can be phonetically transcribed as: [bɔːl] and [dɔːl]. While not the case in all English accents and dialects, these two words share a vowel in the Canadian English spoken in the Vancouver area and thus form a minimal pair.

5 This stimulus can be phonetically transcribed as: [niːm].
The two objects presented during the habituation and test phases were a green and red ball and a doll with light blue clothing and bright yellow hair. These objects are not only distinct in color and shape, but are also highly representative instances of their respective categories (see Figures 1a and b). A multicoloured toy water wheel (“spinner”) was used for both the pre- and post-tests (see Fig. 1c). The first author took individual digital pictures of the ball and doll objects on a black background. These digital pictures were animated to move back and forth across the screen at a slow and constant velocity (“ball” = 12.1 cm/s; “doll” = 12.6 cm/s). Importantly, change in direction of movement was out of synchrony with presentation of the word, to ensure that the infant had no assistance from modal or causal cues (see Gogate & Bahrick, 1998). The “spinner” object was videotaped against a black background and then transferred to laser disc format. The “spinner” was filmed with the base remaining stationary while the wheel was moved around in a clockwise motion. At the distance tested, the objects take up a 13.5° vertical and 13° horizontal visual angle.

**Figure 1**
Objects used during habituation and testing: (a) Ball, (b) Doll, and (c) Spinner

2.3 **Apparatus**
Testing took place in an 2.8 m by 2.3 m quiet room, which was dimly lit by a shaded 60W lamp situated 80 cm to the left of the infant at a 45° forward angle. The infant sat on
the parent’s lap facing a 27 inch Mitsubishi CS-27205C video monitor that was approximately 1.2 m from the infant. The audio stimuli were delivered at 65 dB, ± 5 dB, over a BOSE 101 speaker, located directly above the monitor. The monitor was surrounded by black cloth, which stretched the width and height of the room. The infants were recorded using a Sony DCR-TRV11 digital video camera. The lens of the digital video camera peeked out of a 6.4 cm hole in the black cloth located 21 cm below the monitor. As a masking control during testing, the parent wore Koss TD/65 headphones over which female vocal music was played from a Sony CFD-V17 CD player.

Habit 2000, a computer program produced by Leslie Cohen’s laboratory at the University of Texas at Austin, was used to order stimuli presentation and collect looking time data. The program was run on a Macintosh Power PC G4. Both the visual stimuli and audio stimuli were played from digitized files on the computer and were sent to the monitor and speaker in the testing room.

The experimenter, who was blind to the audio stimuli being presented and to whether a trial was a habituation or test trial, monitored the infant’s looking times via a closed circuit television system from an adjacent testing room. A designated key was pressed on the computer keyboard during infant looks, which the Habit 2000 program recorded. The video record was used for subsequent reliability coding.

2.4 Procedure

After the procedure was explained to the parent or parents and they had signed a consent form, the experimenter obtained the MacArthur Communicative Development Inventory, a vocabulary checklist, from the parent(s) to ensure that the infant comprehended one or both of the target words. The infant and one parent were then taken to the testing room and positioned for the experiment. The experimenter turned on the digital video camera and entered the adjacent observation room to begin testing. The infant was assigned to participate in a preselected order, chosen from a randomly sequenced list of possible orders. These orders counterbalanced the order of test trial (“same” before “switch”/“switch” before “same”) and the type of switch between the test trials (switch in object/switch in word).

The infants were tested using a modified habituation paradigm, identical in structure to that used by Werker et al. (2002), but modified for habituation criterion (adjusted from 50% of the highest total looking time — summed across a block of four trials — to 65% in the current study in order to make it comparable to Stager & Werker, 1997). Each trial began when the infant fixated on a flashing red light. On the first trial, infants were presented with a pretest stimulus, the label “neem” paired with the spinner. During the habituation phase the infant was shown two word object pairs (e.g., Pair A: word ‘ball’ and ball object, Pair B: word ‘doll’ and doll object). Every block of four trials contained two instances of each word-object pairing presented in a random order (ABAB, ABBA, etc.). Looking time was calculated on-line, and when the average looking time across a four-trial block decreased to the preset criterion, the habituation phase ended. The infants participated in a minimum of eight and a maximum of 24 habituation trials.

Following habituation, the test phase began. One test trial was a “same” trial
in which one of the pairings presented in the habituation phase was presented again (e.g., pair A). The other trial, the “switch” trial, contained a familiar word and familiar object but in a novel pairing (e.g., label from pair A with object from pair B). The order of presentation of the trials was counterbalanced across subjects. It was expected that, if infants had accessed the phonetic detail, they would detect the “switch” and look longer during the “switch” than the “same” trial. In the final post-test trial, the child was again presented with “neem” and the spinner. It was expected that if infants were still involved in the experiment, looking time would recover to near pretest level during this final trial.

### 2.5 Reliability coding

To determine the reliability of the experimenter’s coding, a second trained coder scored the looking times of 25% of the usable subjects off-line. On-line scores were rounded to the nearest 0.1s. Off-line scoring was also done to the nearest 0.1s. A Pearson product-moment correlation of on- and off-line scores had to be equal to or greater than .95 for the data to be considered reliable. This level of agreement was reached for all subjects.

### 3 Results

To determine whether infants maintained interest throughout the experiment and recovered from habituation, a series of planned orthogonal comparisons were run to first compare pretest to post-test and, if these two trials were found to be the same, to then compare these trials to the last habituation block. There was no significant difference between the pretest and post-test. The pretest and post-test were significantly different from the last habituation block, $t(2, 39) = 15.61, p < .001$; Mean$_{PRETEST} = 18.7$, Mean$_{LASTBLOCK} = 6.6$, Mean$_{POSTTEST} = 16.6$. A 2 (sex: female vs. male) × 2 (trial block: first four habituation trials vs. last 4 habituation trials) mixed ANOVA produced a significant main effect for trial block, with neither a main effect for gender nor an interaction, $F(1, 14) = 538.23, p < .001$; Mean$_{FIRSTBLOCK} = 14.4$, Mean$_{LASTBLOCK} = 6.6$. Thus, as expected, there was a significant drop in looking time across the habituation phase.

The main set of analyses addressed infants’ performance on the test trials. A 2 (sex: female vs. male) × 2 (test trials: same vs. switch) mixed ANOVA revealed a significant main effect for test trials, with the infants looking longer to the switch trial than to the same trial, $F(1, 12) = 22.48, p < .001$; Mean$_{SWITCH} = 8.49$, Mean$_{SAME} = 4.95$. There was no main effect for sex and no interaction. All infants habituated to the stimuli and all but one of the 16 infants looked longer to the “switch” trial. This shows that infants of 14-months notice the switch in well-known minimal pair word-object pairings (see Fig. 2).

---

6 The variances of the dependent variables lacked homogeneity. The t-test was corrected to account for this nonhomogeneity.
Discussion

In comparing the two studies to date that have demonstrated that 14-month-old infants can attend to fine phonetic detail in words paired with objects (i.e., the current study and Swingley & Aslin, 2002), we see one major commonality. That commonality lies in the use of already known words. In both studies, the infants were tested on highly familiar object-label combinations, rather than the novel object-label combinations used in our previous research. Even though different tasks were used in the two studies, the infants showed success in distinguishing phonetically similar words. Importantly, the infants in the current study succeeded in the identical Switch procedure as used in our previous work. This provides an unequivocal demonstration that lack of attention to fine phonetic detail, a robust finding in word learning situations, does not characterize performance in word recognition tasks. It is noteworthy that the infants in the current study accurately attended to the exact same phonetic detail (the place of articulation cues in [b] and in [d]) that infants in our previous studies failed to access when the task required word learning (Pater, Stager, and Werker, under review; Stager & Werker, 1997; Werker et al., 2002). It would seem that the infants’ a priori knowledge of the words and their referents allowed them to access the relevant phonetic detail and thus notice a violation in the object-label pairing.7

---

7 It is important to note that we use the term knowledge, rather than experience. It is not that the infants have previous experience with the specific object-label combinations presented in our study. None of the infants would have previously heard our speaker, a graduate student in our department, utter “doll” or “ball.” Also, it is very doubtful that the participants would have seen the specific object used in the study. Therefore, the infants were applying their generalized knowledge of the objects and labels used in the study.
4.1 Resource limitation

Our demonstration of novice word learners’ success in distinguishing minimal pair words, together with the previous evidence from Swingley and Aslin’s (2002) mispronunciation study and even the older evidence from Barton (1978, 1980), provides compelling evidence that when infants know words well, fully specified representations are available. As reasoned in the introduction, a resource limitation hypothesis can account for this intriguing pattern of findings. The use of well-known words reduces the demands placed on the infant since the basic object-label mappings of these words have been previously established. Without having to devote cognitive resources to forging the object-label link, novice word learners are better able to access the fine phonetic detail present in the words.

The limited resource explanation requires consideration of other situations that might make access to the fine phonetic detail possible. One might suggest, at first blush, that more frequent exposure to the word-object link would facilitate access to the criterial information. However, we think this is unlikely because in our previous word-learning studies infants were presented with up to 120 repetitions of each word-object pairing (see also Swingley & Aslin, 2002, for a similar argument). We suggest instead it is the conditions under which the child learned the word-object pairing that are important. In every situation in which the children at 14-months succeeded, they came into the testing situation already knowing the word-object pairings. In the course of learning these pairings in their home environments, the infants were likely exposed to each word not only a number of times, but across a number of different speakers in a number of different contexts. Similarly, they were likely exposed to multiple instances of the target objects ball and doll. We would argue that variability in exposure may enable the infants to recognize the new instances of the objects as familiar, and may thus facilitate attentional focus to what is constant in the word forms that are linked to these objects. This line of reasoning would predict that infants of 14-months might fare better in the minimal pair Switch task if the habituation phase included more variability (for an example of this in another word learning task see Hollich, Jusczyk, & Brent, 2002; see also Singh, Bortfeld, & Morgan, 2002).

We have offered an explanation for why infants of 14-months can access the phonetic detail in well-known words but not in a word learning situation. The question still remains unanswered as to what happens between the ages of 14 and 17 months that allows the older infants to access [b] and [d] in new words (see Table 1, p.259). Infants of 17-months have a significantly larger vocabulary than do infants of 14-months. This likely includes not only the words “ball” and “doll,” but many other words beginning with [b] and [d] as well. With a sufficient (and or dense) vocabulary, an infant may be able to pull out the phonemic similarities among these words (see also Beckman

---

8 It could be argued, based on our data, that perhaps infants only have access to detail in well-known minimal pair words and our claim that infants have access to the representations of all well-known words is too broad. Evidence against this position is provided in a study by Swingley and Aslin (2002). They demonstrated that infants of 14-months access phonetic detail in familiar words for which they had close neighbors in their lexicon (e.g., ball) and words for which they had no neighbors (e.g., baby).
& Edwards, 2000; Plunkett, Sinha, Moller, & Strandsbury, 1992; Schafer & Mareschal, 2001). This can then serve to direct attention to the criterial detail in the word learning task even in the laboratory where the words are all pronounced by the same speaker with the same intonation (see Werker & Curtin, in preparation). This explanation predicts the correlation seen between vocabulary size and success in the minimal pair word learning Switch task in infants of 14-months of age in Werker et al. (2002).

A resource limitation hypothesis that involves a single, detailed representation can account for all the results in Table 1. The younger infants (7.5- and 8-month-old infants) can access the phonetic detail in the representation because they are not trying to connect the word forms to objects or concepts. At the end of the first year of life, infants are beginning to link words with their referents and do not seem to utilize all the detail in the word form representation due to the cognitive load of the linking process. However, they can access the detail in the word form representation when there is no object present. More importantly, at the same age at which infants fail to access detail in new words, they can access the exact same detail when the link between word and object is already in place. At 17-months of age and beyond, having established multiple instances of word-referent combinations, the infants can more easily pick out the constant parts of the word representation necessary to make an accurate word-referent link. This allows them to recognize the constancy not only in familiar words, but in new words as well.

4.2 An alternate explanation

The resource limitation hypothesis is not the only possible explanation for the data summarized in Table 1. As noted in the Introduction, some researchers have suggested that initial lexical representations are underspecified, and have taken as evidence failures by toddlers and young children to distinguish minimally different words (Brown & Matthews, 1997; Edwards, 1974; Garnica, 1973; Kay-Raining Bird & Chapman, 1998; Merriman & Schuster, 1991; Pollock, 1987; Shvachkin, 1948/1973). We have argued above and previously (see Werker & Fennell, in press) that this is unlikely, and that the failures shown at 14-months reflect difficulty in accessing rather than in representing fine phonetic detail. The failures shown by older infants may reflect another level of difficulty in access to already represented phonetic detail in the face of an even more demanding word comprehension task. Thus, the resource limitation hypothesis allows for continuity in the underlying phonetic representation of words, while at the same time acknowledging that different testing situations may facilitate or hinder access to that phonetic detail.

Admittedly, it is still an open question as to whether lexical representations are fully specified at the time of attempting to learn a new word, or only become fully specified after the word is well known. We would argue, however, that the most parsimonious explanation is that the word is fully specified from the beginning and that the difficulty lies in access rather than in the representation. As outlined in Werker & Curtin (in preparation), one way to envision why 14-month-old infants fail to attend to detail in our word learning task is to imagine “bih” and “dih” represented as fully specified word forms that overlap on all dimensions except place of articulation.
<table>
<thead>
<tr>
<th>Authors</th>
<th>Age in months</th>
<th>Procedure</th>
<th>Main Difference, aside from age, from Stager &amp; Werker, 1997 (Exp. 1)</th>
<th>Contrast Used</th>
<th>Did the Infants Use Phonetic Detail?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jusczyk &amp; Aslin, 1995</td>
<td>7.5</td>
<td>Head-turn Preference</td>
<td>Different testing procedure</td>
<td>[c]–[t], [b]–[d]</td>
<td>Yes</td>
</tr>
<tr>
<td>Stager &amp; Werker, 1997</td>
<td>8</td>
<td>Switch</td>
<td>Simplified version of the Switch Procedure</td>
<td>[b]–[d]</td>
<td>Yes</td>
</tr>
<tr>
<td>Hallé and de Boysson-Bardies,1996</td>
<td>11</td>
<td>Head-turn Preference</td>
<td>Use of familiar word forms; different testing procedure</td>
<td>Multiple</td>
<td>No</td>
</tr>
<tr>
<td>Stager &amp; Werker, 1997 (Exp. 1)</td>
<td>14</td>
<td>Switch</td>
<td>**</td>
<td>[b]–[d]</td>
<td>No</td>
</tr>
<tr>
<td>Stager &amp; Werker, 1997 (Exp. 2b)</td>
<td>14</td>
<td>Switch</td>
<td>Simplified version of the Switch Procedure</td>
<td>[b]–[d]</td>
<td>No</td>
</tr>
<tr>
<td>Pater, Stager, &amp; Werker, under review</td>
<td>14</td>
<td>Switch</td>
<td>Use of well-formed word stimuli (e.g., bin-din).</td>
<td>[b]–[d], [b]–[p], [d]–[p]</td>
<td>No</td>
</tr>
<tr>
<td>Werker, Fennell, Corcoran, &amp; Stager, 2002 (Exp. 2)</td>
<td>14</td>
<td>Switch</td>
<td>More distinct objects; greater exposure to stimuli</td>
<td>[b]–[d]</td>
<td>No</td>
</tr>
<tr>
<td>Stager &amp; Werker, 1997 (Exp. 4)</td>
<td>14</td>
<td>Switch</td>
<td>Labels not associated with objects.</td>
<td>[b]–[d]</td>
<td>Yes</td>
</tr>
<tr>
<td>Swingley &amp; Aslin, 2002</td>
<td>14</td>
<td>Visual Fixation</td>
<td>Use of familiar words and objects; different testing procedure</td>
<td>Multiple</td>
<td>Yes</td>
</tr>
<tr>
<td>Current Study</td>
<td>14</td>
<td>Switch</td>
<td>Use of familiar words and objects</td>
<td>[b]–[d]</td>
<td>Yes</td>
</tr>
<tr>
<td>Werker et al., 2002</td>
<td>17 and 20</td>
<td>Switch</td>
<td>More distinct objects; greater exposure to stimuli</td>
<td>[b]–[d]</td>
<td>Yes</td>
</tr>
<tr>
<td>Swingley &amp; Aslin, 2000</td>
<td>18–24</td>
<td>Visual Fixation</td>
<td>Use of familiar words and objects; different testing procedure</td>
<td>Multiple</td>
<td>Yes</td>
</tr>
</tbody>
</table>
The task facing the child in the Switch task is to take those two fully specified word form representations and attach them to two different objects. Because the two word forms overlap on so many dimensions, considerable attention is required to pull out just that feature — among all the features in the representation — that distinguishes the two words. In the face of the computational difficulty of forging novel word-object links, the cognitive resources are simply not available for such directed attention.

Still, some may reject the parsimony argument as an adequate proof for the resource limitation hypothesis. After all, much of the data in this paper that have been presented in favor of the resource limitation hypothesis could be equally applied to the representation argument. For example, a major claim shared by the two hypotheses involves the developmental data pattern of phonemic perception (outlined in Table 1). The resource limitation explanation for the data has already been presented. The advocates of the representation-based approach would argue that young infants (8-months) and older infants (14-months) listening to labels not matched with objects would access phonetic detail because they are accessing a low-level phonetic representation. The failure of infants aged 14-months to access detail in newly learned words would reflect the inadequate phonological representations that these infants possess due to their incomplete phonemic inventory, which needs to be built up over time (e.g., Brown & Matthews, 1997; Shvachkin, 1948/1973). In this approach, the phonemic inventory is only established as required to distinguish words in the lexicon. Infants of 14-months would not yet know many minimally different words, and would thus not have experienced the pressure to represent fine phonetic detail when learning new words.

One claim made by developmental underspecification theorists is that not all phonemic contrasts are created equal (see discussion in Brown & Matthews, 1997; Rice & Avery, 1995). Phonemic representations are elaborated based on pressure from the linguistic input. Therefore, the infant would begin with no phonemic contrasts and then proceed in a step-wise fashion in establishing a phonemic inventory. Some models predict a fixed order to this progression. Studies with older children have demonstrated a stage-like progression in the comprehension of phonemic contrasts (Barton, 1978, 1980; Brown & Matthews, 1997; Edwards, 1974; Garnica, 1973; Shvachkin, 1948/1973). One interpretation of the resource limitation hypothesis would reason that if the cognitive load preventing access to phonetic detail in the completely specified representation is eased, then all phonemes should be equally accessible. Our interpretation of the resource limitation hypothesis, however, is based on the premise that the infant is listening to the words, but just failing to access some of the information. This allows for the possibility of differential access to more salient acoustic/phonetic detail.

It is also important to note that the studies demonstrating differential use of contrasts have all involved explicit tasks (e.g., picture selection), which are more difficult for young infants and perhaps more unstable. It is still an unanswered question whether or not there is any evidence of an order of phonemic acquisition in more implicit tasks. Indeed, the only study to date explicitly comparing different phonetic contrasts showed that infants aged 14-months fail at place, voicing, and place plus voicing contrasts (Pater, Stager, and Werker, under review).
Since the data presented thus far seems to support both theoretical stances, we must turn to possible ways to disambiguate the competing hypotheses. Our planned study with Daniel Swingley is an important test of the representational underspecification hypothesis. In this study, 14-month-old infants will be habituated to two novel, nonsense minimal pairs (e.g., bin/din). After habituation, the infants will be tested using the Swingley and Aslin (2002) visual fixation task, which may have fewer memory demands than the Switch task since it involves the simultaneous presentation of the habituated object-label pairings at test. If the infants in this proposed study do not access the detail in the habituated words, we would once again have data that supports both hypotheses. However, if the infants of 14-months access the habituated words’ phonetic detail in this study, their success can only be accounted for by the resource limitation hypothesis. The representation hypothesis would be challenged in this case since it posits underspecified representations at this stage of development (Brown & Matthews, 1997; Shvachkin, 1948/1973).

Another way to attempt to disambiguate resource limitation from the representational account is to methodically test infants on their ability to access fine phonetic detail in a well-known word that has no close phonetic neighbor in the infant’s current vocabulary. The resource limitation hypothesis would predict that the infant should notice the difference between the two words since the known object-label pairing would ease the processing load. On the other hand, the representation account would predict that the infant would confuse the known word with the unknown minimally different word, since the known word has no close neighbors in the infant’s vocabulary and thus no need to be fully specified. The mispronunciation task used by Swingley and Aslin (2002) takes this approach. Although we have argued above that the Swingley and Aslin task has different processing demands than the Switch task, their results do suggest that infants can access detail even in words without near neighbors. If these results hold with the Switch task (studies currently underway), then it would lend support to the resource limitation hypothesis.

5 Conclusion

The link between speech perception and word learning has proven to be more complicated than first expected. Still, it would seem that the emerging organization in speech perception seen at the end of the first year does prepare the child for word learning by determining the detail that is available in word forms. Along these lines, we have presented an account that focuses on continuity in the representation, but have argued for differences in access as a function of processing load. We have also presented a competing hypothesis that involves a discontinuity in the underlying representations. Our analysis supports the former account. We believe that this debate is central to clarifying the link between prelexical perception and postlexical phonology. Unpacking

---

9 The representation hypothesis does not necessarily propose that the specific phonemes [b] and [d] are underspecified at this age. However, the representation theorists would have difficulty explaining why same-aged infants would notice the phonetic detail in this task and not in the Switch procedure.
this link through additional tightly controlled, linguistically motivated laboratory studies holds promise for a fuller understanding of phonological development.

**References**


