Stager and Werker (1997) show that 14-month-olds engaged in a word-learning task fail to respond to a switch between the minimal pair [bi] and [di], though they do respond to a switch between [lif] and [nim] in the same task. In this paper we show that the [bi]/[di] results extend to stimuli that respect English phonotactics ([bin] vs. [dun]), to a voicing contrast ([pin] vs. [bin]), and to voicing and place combined ([pin] vs. [din]). Our interpretation of these results is that when a phonological contrast like place or voicing is first acquired, it remains only partially integrated, and can be lost under the processing demands of word learning. We formalize partial integration in terms of unranked Optimality Theoretic constraints, and discuss the predictions of this account for further research.

1. Introduction
Research on phonological acquisition has typically focused on production, examining the development of children’s ability to produce the segmental and prosodic structures of their language. Standard phonological accounts of the shape of children’s productions, from those in Smith (1973) and Stampe (1969), through recent ones framed in Optimality Theory (see Bernhardt and Stemberger 1998; Kager et al. 2004, and the references therein), posit a phonological grammar that generates the child’s output from a lexical representation that relatively closely approximates the adult surface form (cf. e.g. Macken 1980; Velleman 1988). In the earliest stages, the child’s productions are subject to rules or constraints that greatly limit their complexity. For example, when target fricatives surface as stops, as is often the case in early child speech, the contrast between the two categories is eliminated. As acquisition proceeds, the phonological grammar changes to

*Acknowledgments:* Thanks to the parents and children who participated in the studies, to Sharon Lee and Kathleen Corcoran for their help in testing infants, and to Suzanne Curtin, Christopher Fennell, and Doug Pulleyblank for discussion. This research was supported by NSERC Discovery grant RGP-110 to Janet Werker, a Killam doctoral fellowship to Christine Stager, and a SSHRC post-doctoral fellowship to Joe Pater.
allow greater complexity in the child’s productions, so that they come to match the detail in the adult forms.

Relatively little is known about the prior development of children’s ability to perceptually represent phonological structures and encode them in lexical representations. In fact, the fundamental question of whether complexity gradually unfolds in the perceptual domain in ways parallel to production, or whether perceptual and lexical representations are instead unlimited in their complexity from the outset of phonological acquisition, remains largely unanswered. For each of these two positions on the phonological complexity of early representations there is a line of research that seems to support it.

Infant speech perception research has demonstrated the remarkable perceptual capabilities of even the youngest of children (see Best 1994; Jusczyk 1997; Werker and Tees 1999 for recent reviews). In these studies, infants are exposed to a repeated auditory stimulus (e.g. the syllable [ba]), which is then changed along the dimension of interest (e.g. to [da], a change in place of articulation). Sensitivity to this change is measured in a variety of ways, depending on the age of the infant. The most common measure for very young infants is sucking rate and amplitude, while after four to six months gaze and head turns are typically monitored. In their first six months of life, infants have been shown to be able to distinguish nearly every contrast they have been tested on, including ones not present in the ambient language (Eimas et al. 1971; Streeter, 1976). Further, there is evidence of perceptual reorganization in the second half of the first year, in which sensitivity to non-native language contrasts is reduced relative to native ones (e.g. Kuhl et al. 1992; Werker and Tees 1984a). From these demonstrations of the exquisite perceptual abilities of infants, one might conclude that when meaningful words are first acquired in the beginning of a child’s second year of life, lexical representations encode all segmental contrasts.

On the other hand, experiments with slightly older children, at around two to three years of age, document gradual acquisition of segmental contrasts. In minimal pair identification tasks, first used by Shvachkin (1948/1973), a child makes a gestural response to indicate which member of a minimal pair s/he has heard (e.g. by pointing to a ball or a doll when hearing [dal]). Shvachkin found that performance on some contrasts tended to be better than on others, and proposed a universal order of acquisition, similar to that posited for production by Jakobson (1941/1968). Subsequent research using this experimental paradigm has drawn into question the validity of Shvachkin’s proposed order of acquisition (e.g. Barton 1976, 1980; Brown and Matthews 1997; Eilers and Oller 1976; Garnica 1973; Velleman 1988). However, this body of work continues to support
the general view that some contrasts are acquired before others, and that there are therefore restrictions on the complexity of early representations (see esp. Brown and Matthews 1997).

There are many differences between the methodologies of classical infant speech perception research and the minimal pair identification studies, and hence many possible explanations for the differences between the two sets of results. Barton (1976, 1980) uses the terms phonetic and phonemic discrimination to distinguish the infant tasks from the minimal pair tasks; the latter are termed phonemic since they require the subject to pair sound with meaning. A typical infant task looks only for evidence that subjects are sensitive to a change in a repeated auditory stimulus. To succeed on the minimal pair task, however, the subject must be able not only to perceive an acoustic distinction, but also to correlate that distinction to a difference in meaning. In terms of standard linguistic levels of representation, we might say that infant tasks tap phonetic representations, while minimal pair tasks access phonological representations.

Under that view, then, we would say that the minimal pair tasks show that receptive phonological competence does unfold in a way parallel to production, and that the perceptual abilities evidenced in the first year of life are phonetic, rather than phonological. However, Barton (1976, 1980) also draws attention to the considerable task demands imposed by the minimal pair methodology, suggesting that these might mask infants’ ability to lexically encode contrasts. In requiring the child to make a gestural response to indicate which member of a minimal pair was heard, these experiments impose cognitive demands that may well lead to an underestimation of a child’s ability to lexically encode a contrast. It is possible that two-year-old children have phonologically detailed lexical representations that are simply not accessed by the minimal pair task.

In this paper, we discuss recent research that bridges the gap between these two sets of results by employing methodologies that add the dimension of meaning to infant speech perception, while reducing task demands relative to minimal pair experiments. The reduction of task demands has the added benefit of allowing the same methodology to be used across a range of ages, thus forming another bridge between the two earlier lines of work. Section 2 discusses research using the ‘Switch’ technique of Werker et al. (1998), in which looking time is the dependent variable in measuring children’s ability to form sound-meaning pairings. This research indicates that at about 14 months of age, infants ignore a difference in consonantal place of articulation ([b] vs. [d]) when learning new word-object pairings, even though they are able to perceive this distinction (Stager and Werker 1997). By about 17 months, or more precisely when their vocabulary reaches a sufficient size, infants in this same task are able to associate a place of
articulation difference with a difference in meaning (Werker et al. 2002). In sections 3 through 5 we present new research that shows that 14-month-olds’ failure to respond to a switch is also observed when the words respect the English phonotactic restriction against word-final lax vowels ([bn] vs. [dm]), and that this failure extends to a voicing contrast ([pin] vs. [bn]), and to voicing and place combined ([pm] vs. [dm]).

This set of results appears to indicate that when phonological representations are first acquired receptively, they are indeed reduced in complexity relative to the adult form, encoding neither place nor voicing contrasts. However, in section 6 of this paper, we discuss further results that show that when well-known words are used, 14-month-olds do respond to the place contrast in the ‘Switch’ task (Fennell and Werker 2003), and that they also distinguish minimal pairs in a visual fixation procedure (Swingley and Aslin 2002). In these experiments, the infants are engaged in word recognition; in those of Stager and Werker (1997) and the present paper, the infants are engaged in word learning. Here, levels of representation are not at issue: both tasks are clearly phonological, in that they require the infant to pair a meaning distinction with a sound distinction.

One way of understanding the discrepancy between these results is that when a contrast is first acquired, it is not stable, and can be lost under processing demands; word learning is clearly more demanding than recognition (Fennell and Werker 2003; Werker and Curtin in prep.). This is parallel to production, in which a new structure occurs only variably when it is first acquired. Drawing on proposals in Optimality Theory for accounting for variation in production, we provide an account of the partial receptive acquisition of a contrast in terms of constraints whose ranking is not fixed, and whose ordering can be affected by processing demands. We then discuss an extension of this account to other results, including those of the classical minimal pair task, and outline some predictions for further research.

2. Assessing sound and meaning pairing
One might take the development of language-specific perception in the second half of the first year as evidence of the emergence of a phonological system. However, having receptive command of a phonological contrast involves not only perceiving it, but also storing it in long-term memory, and using it in the service of a meaning distinction.

A study by Jusczyk and Aslin (1995) shows that at 7.5 months, infants are capable of representing place of articulation contrasts in memory. The aim of the study was to determine whether infants are able to recognize words in a fluent speech context. They used a variant of the headturn preference procedure that has a familiarization and a test
phase. In the first phase, infants are played CVC words pronounced in isolation, and are thus familiarized with the words. At the start of a trial, a light flashes above one of the speakers mounted on either side of the testing booth, and the stimuli are played so long as the infant looks that speaker. The familiarization session lasts until a total of 30 seconds listening time is accumulated for each word. In the test phase, a passage of running speech is played from a speaker, and the amount of time that the infants orient toward the speaker is measured. Listening time (i.e. orientation time) is compared for two types of passages: ones that contained the words from the familiarization phase, and similar passages that that did not contain the familiarized words. Jusczyk and Aslin found that infants listened longer to the passages that contained words from the familiarization procedure. What is important in the present context is that this result did not obtain when the familiarized words differed from the test words in the place of articulation of the initial consonant. When familiarized with isolated words like *bog* and ‘*tup*’ infants did not show a preference for passages containing *dog* and *cup* over passages that lacked the minimal pairs. The difference between this condition, and the one in which they were familiarized with *dog* and *cup* and did show a preference for the previously heard word, shows that infants are capable of retaining the place contrast in memory, at least between the familiarization and test phases.

As Jusczyk and Aslin make clear, these results show that infants are capable of representing word forms in considerable phonetic detail, but do not indicate whether infants are attaching any meaning to the words. As such, they do not allow us to establish whether a place of articulation contrast can be used to support a meaning difference. Indeed, work by Hallé and de Boysson-Bardies (1996) would suggest that meaning interferes with access to phonetic detail. When tested on familiar words overheard in everyday speech, infants of 11-12 months listened longer to these items over unfamiliar foils. However, when the unfamiliar foils differed in only a single phonetic feature from the familiar words, the infants appeared unable to distinguish them. The authors raise the possibility that the older infants may have had difficulty because they, unlike the younger infants in the Jusczyk and Aslin work, were listening for meaning.

This question is examined by Stager and Werker (1997), who adapted a technique developed by Werker *et al.* (1998) in which looking time is the dependent variable in measuring children’s ability to form sound-meaning pairings (cf. Schafer and Plunkett, 1998; Swingley and Aslin 2000, 2002 for related but different approaches). This methodology reduces task demands considerably from the minimal pair experiments of Shvachkin (1948/1973) *et seq.*, in which children are required to gesturally indicate which of two objects corresponds to the word they have heard. In the ‘Switch’
methodology, infants are familiarized with two repeated novel syllables, each of which is paired with a brightly colored moving object on a video display. Each of the sound-object pairings is repeatedly presented in separate trials lasting from 14 to 20 seconds, depending on the particular design. This exposure continues until the infant is habituated, which is determined by a decrease in looking time. Following habituation, test trials occur in which a sound-object pairing has been changed (switch trials), and in which the sound-object pairing is the same as in familiarization (same trials). If infants have learned the sound-meaning associations, they should look longer during the switch than the same trials, since the switched pairing will be new to them, and this novelty will cause dishabituation.

Werker et al. (1998) show that at 14 months (but not 8 to 12), infants are able to form sound-meaning pairings when the objects are moving, and when the labels are highly dissimilar ([lff] vs. [nim]). Stager and Werker (1997) tested whether 14-month-olds would also notice a switch in sound-meaning pairing when the words differed only in the place of articulation of the initial consonant ([bi] vs. [di]). Unlike the Werker et al. (1998) findings for [lff]/[nim], 14-month-olds did not look significantly longer during the switch than the same trials. Furthermore, when the task was simplified by including only a single sound-object pair during familiarization, and switching the sound in the switch trials, the 14-month-olds continued to show no evidence of noticing the change in place of articulation. But as with the two pair familiarization condition, when the contrast was between the more dissimilar [lff] and [nim], the 14-month-olds did look significantly longer during the switch than during the same trials.

Interestingly, when 8-month-olds were familiarized with a single sound-object pair condition and tested on a switch between [bi] and [di], they did look significantly longer during the switch trials. Stager and Werker (1997) interpret this as indicating that the infants at these two ages were treating the task differently: For the 8-month-olds, it was being performed as a simple phonetic discrimination task, whereas for the 14-month-olds, it was being treated as a word-learning task. This interpretation is bolstered by two other findings. First, Werker et al. (1998) found no evidence of 8-month-olds being able to form a sound-object association in the two-pair condition, even when the sounds were the very different [lff] and [nim]. And second, Stager and Werker (1997) present evidence that the 14-month-olds were able to notice the switch when the task was changed by removing the element of word learning. To do so, they replaced the brightly colored moving object on the video display with an unbounded checkerboard pattern, which is unlikely to be perceived as a nameable object (Spelke, Vishton, and Van
Hofsten 1994; Woodward 1993). In this condition, the 14-month-olds did listen longer during the switch trials.

This set of results is summarized in the following table:

| Age      | Experiment Type                              | Switch noticed?
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>8 months</td>
<td>2 sound-object pairings ([lɪf]/[nim])</td>
<td>No</td>
</tr>
<tr>
<td>8 months</td>
<td>1 sound-object pairing ([bɪ]/[dɪ])</td>
<td>Yes</td>
</tr>
<tr>
<td>14 months</td>
<td>2 sound-object pairings ([lɪf]/[nim])</td>
<td>Yes</td>
</tr>
<tr>
<td>14 months</td>
<td>2 sound-object pairings ([bɪ]/[dɪ])</td>
<td>No</td>
</tr>
<tr>
<td>14 months</td>
<td>1 sound-object pairing ([bɪ]/[dɪ])</td>
<td>No</td>
</tr>
<tr>
<td>14 months</td>
<td>1 sound-object pairing ([lɪf]/[nim])</td>
<td>Yes</td>
</tr>
<tr>
<td>14 months</td>
<td>1 sound-checkerboard pairing ([bɪ]/[dɪ])</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 1: Results from Stager and Werker (1997)

Thus, when first learning a word, 14-month-olds treat words that differ in the place of articulation of initial consonants as identical, even though they are able to distinguish them in a perceptual task. This suggests that the contrast has not been fully integrated into their phonology; how this partial integration might be formalized is discussed in the final section of this paper.

The question of when the place of articulation distinction is made use of in word learning is addressed in Werker et al. (2002). They used the two pair version of the Switch methodology, with [bɪ] and [dɪ] as stimuli, in a study with 14-month-old, 17-month-old, and 20-month-old subjects. The two pair methodology was employed because it is unambiguously a word learning task, in contrast to the one pair methodology, which may be passed as a phonetic discrimination task. They found that at 17 and 20 months, but not at 14 months, looking time was longer in the switch trials.

Having gathered parental reports of vocabulary size, Werker et al. (2002) were able to assess whether there was a correlation between vocabulary size and performance on the Switch task. They found that at 14 months, there was a significant correlation between size of productive vocabulary and performance on the Switch task, with the correlation between comprehension vocabulary and performance being significant in one analysis, and nearly so in another. At 17 months, the vocabulary size/task performance correlation was nearly significant for comprehension vocabulary, and positive, but non-significant, for production vocabulary. Further post-hoc analyses suggest that, with this group of infants, the vocabulary threshold at which infants begin to look significantly
longer during the switch trials was 25 reported words in production, and 200 reported words in comprehension.

A potential concern does arise from the fact that the forms [bi] and [di] used in these experiments are actually ill-formed as English words. Monosyllabic content words are always made up of either a tense vowel (e.g. [bij], [bej]) or a lax vowel followed by a consonant (e.g. [btt], [bet]). Monosyllables headed by lax vowels without a following coda consonant (e.g. [bi], [be]) do not occur as English content words. A series of studies have shown adult sensitivity to this constraint, Moreton (1999) demonstrates that English listeners are more likely to identify a vowel that is ambiguous between [ij] and [i] as [ij] in the word-final context than in a context where both are permitted. Cebrian (2002) shows that native English speakers, and Catalan learners of English, use it in interpreting the morphological composition of nonce words. Finally, Pater and Tessier (2003) show that the restriction against words of this type plays a role in English speakers’ acquisition of phonological alternations in artificial languages.

One might therefore worry that the sub-minimal status of [bi] and [di] is a hindrance for the establishment of lexical representations for these words. To address this issue we employed as stimuli [bm] and [dn] instead of [bi] and [di]. Because of the density of the lexicon of monosyllabic words, it was impossible for us to use non-words that would differ minimally along the dimensions of interest. However, we did choose words that would be unlikely to be familiar to infants of 14-months. We used a parental questionnaire to determine whether any of the words were familiar to the infants; if they were, the subject was not included in analyses.

A further issue is that the experiments reported in Stager and Werker (1997) and Werker et al. (2002) examine only the place of articulation contrast ([bi]/[di]). To assess whether the difficulty that 14-month-olds had with minimally contrastive words would generalise beyond place of articulation, we also ran the word learning experiment with pairs differing in voicing ([bm]/[pm]), as well as a combination of place and voicing ([pm] vs. [dm]). The results of these three new experiments are presented in the following three sections.

3. Experiment 1: Word form

In the studies reported by Stager and Werker (1997) and Werker et al. (2002), 14-month-old infants could have failed to learn the words [bi] and [di] not because they have difficulty encoding and representing the consonantal place contrast, but because they have difficulty with words that disobey the English phonotactic restriction that words must be minimally bimoraic. This experiment was designed to assess this possibility by
testing infants aged 14-months on the same place of articulation distinction, but this time embedded in proper bimoraic CVC word forms: [bɪn] vs. [dɪn].

3.1 Participants
Sixteen 14-month-old infants completed this study, 8 girls and 8 boys (mean age, 14 months 12 days; range, 14 months 1 day to 14 months 21 days). All subjects were without apparent health problems and were exposed to English at least 90% of the time, as assessed by parental report. An additional 10 infants were tested but were not included in the analyses because they were too fussy to continue (n = 7), their parents interfered in some way (n = 1), or because of experimenter error (n = 1) or equipment failures (n = 1).

Infants were recruited mainly through visiting new mothers at BC Women's Hospital, but also through voluntary response to public service announcements. At the time of recruitment, parents consented to be contacted about the possibility of participation. Participating infants were given an “Infant Scientist” t-shirt and diploma.

3.2 Stimuli
The audio stimuli were infant-directed repetitions of the CVC words [bɪn] and [dɪn]. Infant-directed speech (IDS) was used as it has been shown to be effective in gaining and maintaining infant attention (Fernald, 1985; Werker and McLeod, 1989) and has been shown to facilitate word learning in infants (Fernald, McRoberts, and Herrara, 1991). The use of IDS has also been shown to simplify discrimination for the infant (Karzon, 1985). The test stimuli [bɪn] and [dɪn] differ in only the place of articulation of the initial consonant. An additional nonsense label, [pʰɔk], was presented during the pre- and post-test trials. The stimulus [pʰɔk] was chosen as it is different in consonant and vowel features from the target nonce words.

A female speaker of Midwestern American English recorded the tokens in a soundproof booth. She produced several exemplars of each syllable in an infant-directed, rise-fall intonational phrase. Final stimuli comprised seven exemplars of approximately 0.7 s in duration each, with a 1.5 s silent interval between exemplars, resulting in audio files of 14 s in duration. The seven exemplars were ordered to maintain the rise-fall intonation pattern.

The video stimuli were two objects ("crown" and "roundy"), both made with yellow, blue, and red modelling clay, but differing in shape. These stimuli were used for the habituation and test trials. A store-bought, multicoloured toy water wheel ("spinner") was used for both the pre- and post-tests. All three objects were videotaped against a black background and then transferred to laser disk format. “Crown” and “roundy” were
taped moving back and forth across the screen at a slow and constant velocity. The “spinner” was filmed with the base remaining stationary while the wheel was moved around in a clockwise motion.

3.3 Equipment and Apparatus

The experiment took place in a dimly lit quiet room. The infant sat on the parent’s lap facing a video monitor, situated at a distance of approximately 70 cm from the infant. The audio stimuli were delivered from a speaker, located directly above the monitor. The monitor was surrounded by black cloth, which stretched the width and height of the room. The infant’s looking times were recorded using a video camera whose lens peeped out of a 6 cm hole in the black cloth located 25 cm below the monitor.

To prevent the parent from hearing the stimuli and inadvertently influencing the child during testing, the parent wore headphones over which female vocal music was played. The experiment was controlled by a version of the Habit program, created by the Leslie Cohen laboratory at the University of Texas at Austin. The experimenter was blind to the audio stimuli presented, and to whether the trial was a habituation or test trial. The experimenter monitored the infant’s looking times via a closed circuit television system in the observation room. A designated key was pressed on the computer keyboard during infant looks, which the Habit program recorded. The video record was used for subsequent reliability coding.

For further details on the experimental apparatus, see Werker et al. (1998).

3.4 Procedure

The parent or parents who brought the child to the lab were queried to find out whether their children might have been exposed to the word bin (some parents of British origin refer to the trashcan as the trash bin) or the word din. Any infant who was exposed to either of those words was excluded from the study.

The infant was assigned to participate in a pre-selected order, chosen from a randomly sequenced list of possible orders. One male and one female infant were assigned to each of the 8 possible orders. These orders counterbalanced the order of test trial ("same" before "switch"/"switch" before "same").

The infants were tested using a modified habituation paradigm, identical in structure to that used by Stager and Werker (1997; Experiment 1; see also Werker et al., 1998). When the infant fixated to a flashing red light, the trial began and lasted for 14 s. On the first trial, infants were presented with the pre-test stimulus, the label [pʰk] paired with the spinner. During the habituation phase the infant was familiarized with two word-
object pairs, each presented in separate trials (e.g., Pair A: word [dɪn] and object "roundy", Pair B: word [bɪn] and object "crown"). 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 by the custom software, and when the average looking time across a block of two-trials decreased to a preset criterion (65% or less of the first block of two trials), the habituation phase ended. The infants were given a minimum of 8 and a maximum of 28 trials to reach criterion. Following habituation, infants were presented with two test trials: a Same trial which was identical to one of the trials in the habituation phase (e.g. [dɪn] with "roundy" was again presented), and a Switch trial in which the familiar object and familiar word were presented in a new combination, (e.g. [dɪn] with "crown"). Order of Same and Switch trials were counterbalanced across infants. Following the two test trials, a post-test trial was presented in which the spinner and the word [pʰɔk] were again presented.

3.5 Results
To determine the reliability of the experimenter’s on-line coding, a second trained coder rescored the looking times of 25% of the useable subjects off-line. A Pearson product-moment correlation was used to estimate the amount of agreement between the two coders. Perfect agreement would result in a correlation of 100%. We required a correlation of equal to or greater than 95% for the data to be considered reliable. This level of agreement was reached for all subjects.

To ensure that infants maintained interest throughout the entire experiment and did not become fatigued, a preliminary analysis was conducted comparing infant looking times on the pretest and posttest trials. The ANOVA used pretest and posttest looking times as the repeated measures, within group factor. In this and all subsequent analyses, we also included male vs. female as between group factors to allow exploration of gender differences in performance, since in previous work we have sometimes found successful performance by females but not males (e.g. Werker et. al. 1998). A 2 X 2 mixed ANOVA (female vs. male, pretest vs. posttest) yielded no main effects or an interaction. Thus, there was no evidence that the infants fatigued during the course of the experiment.

To determine whether the infants did habituate we compared looking time in the first two and last two trials of the habituation phase of the experiment. A 2 X 2 mixed ANOVA (female vs. male, first two vs. last two habituation trials) revealed a main effect for trial block \( F(1, 14) = 193.361, p < .0001 \) Mean
\(_{FIRSTBLOCk} = 12.03 \) s, Mean
\(_{LASTBLOCk} = 13.43 \) s, \( b = 0.44 \), \( MSE = 7.89 \) s. Mean
\(_{FIRSTBLOCk} = 12.03 \) s, Mean
\(_{LASTBLOCk} = 13.43 \) s, \( b = 0.44 \), \( MSE = 7.89 \) s.
= 5.37 s]. There was no main effect for sex and no interaction. These results indicate that subjects did indeed habituate during the course of the experiment.

The main set of analyses addressed infants’ performance on the test trials. If the infants were able to learn the sound-object associations and to discriminate a change to a new CVC differing only in place-of-articulation, we would expect longer looking times to the “switch” trial than to the “same” trial. A 2 X 2 mixed ANOVA (female vs. male, same vs. switch) revealed no significant main effects or interactions (Mean_{SWITCH} = 7.49 s, Mean_{SAME} = 7.12 s). Thus, the infants familiarized to [bin] and [din] each paired with one of the objects “roundy” and “crown”, and then tested with a switched pairing, failed to show a recovery in looking time. This indicates that they did not notice the switch to the minimally different label.

3.6 Discussion

As discussed above, the fact that the stimuli used in the studies reported by Stager and Werker (1997) and Werker et al. (2002) did not conform to a phonotactic restriction in English against monosyllables ending in lax vowels allows for an alternative explanation for the finding that infants did not notice a switch in place of articulation in the word learning task. The results of the present study argue against that alternative, since even when the words included a word-final consonant, and hence conformed to English phonotactics, 14-month-old infants still failed to notice a switch in the place of articulation of the initial consonant. Therefore, we can conclude that the infants had difficulty with these words not because they violated English phonotactics, but because they differed in only consonantal place of articulation.

4 Experiment 2: Perception of Voicing Switch in Word Learning

This experiment was designed to see if the results with place of articulation would generalize to another contrast: Do 14-month-old infants also fail to respond to a switch in consonantal voice when involved in a word learning task? In addition, we used this experiment as an opportunity to determine whether using the single pair methodology with stimuli meeting English phonotactics would result in infants noticing the switch between aural stimuli. Stager and Werker (1997) show that 14-month-olds fail to respond to the place of articulation switch in both the single word-object familiarization condition (Experiment 2) and the two word-object familiarization condition (Experiment 1). However, since the single pair methodology may be less difficult, we decided to see whether this potentially more sensitive task would reveal evidence of the use of a consonantal contrast in word learning.
4.1 Method
The Method was similar to that used in Experiment 1, but this time infants were taught only a single word-object association and were tested on their ability to detect a change to a minimally different word.

4.2 Participants
Sixteen 14-month-old infants completed this study, 8 girls and 8 boys (mean age, 14 months 14 days; range, 13 months 29 day to 14 months 26 days). All subjects were without apparent health problems and were exposed to English at least 90% of the time. An additional 18 infants were tested but were not included in the analyses because they were too fussy to continue (n = 7), their parents interfered in some way (n = 4), they were reported to know one of the words in question (n = 2) or because of experimenter error (n = 5). Subjects were recruited as in Experiment 1.

4.3 Stimuli
The audio stimuli were as in experiment 1, except this time the CVC syllables were [b\textit{in}] and [\textit{p}h\textit{in}]. These syllables differ in only the voicing of the initial consonant; (partially) voiced, unaspirated [b] vs. voiceless, aspirated [\textit{p}h]. The nonsense label, [\textit{p}h\textit{k}], was again presented during the pre- and post-test trials. These stimuli were recorded and prepared as in Experiment 1.

The video stimuli differed from Experiment 1 in that only a single object, “roundy” was used during the habituation phase and test trials (as in Stager and Werker 1997; Experiments 2 and 3). The video stimulus used during the pre- and post-test trials was again the “spinner”.

4.4 Equipment, Apparatus, and Procedure
The equipment and apparatus were identical to those used in Experiment 1. The experimental procedure was as in Experiment 1, except for the following modifications. Parents were again queried to see if they believed the infants could have been exposed to the word \textit{bin} but instead of \textit{din} they were asked about \textit{pin}. Any infant who was exposed to either of those words was excluded from the study.

In this experiment infants were tested using a modified habituation paradigm, identical in structure to that used by Stager and Werker (1997; Experiments 2 and 3). The pre-test was as in Experiment 1. During the habituation phase the infants were shown the object "roundy" paired with either the label [\textit{b}in] (for half the infants) or the label [\textit{p}h\textit{in}]. Following habituation, infants were presented with two test trials: a Same trial in which
roundy and familiar CVC (e.g. [bɪn]) were again presented, and a Switch trial in which roundy was presented again, but this time with the minimally different CVC (e.g. [pʰɪn]). Order of Same and Switch trials was counterbalanced. The post-test trial was as in Experiment 1.

4.5 Results
The reliability of the experimenter’s coding was assessed as in Experiment 1. A Pearson product-moment correlation of equal to or greater than 95% was reached for all subjects whose looking time was rescored, thus indicating that the measurement of looking time was reliable.

Also as in Experiment 1, we checked for fatigue during the course of the experiment by performing a 2 X 2 mixed ANOVA (female vs. male, pretest vs. posttest). This yielded no main effects or an interaction, thus providing no evidence of fatigue.

To ensure that the infants did habituate, we examined infant looking times during the habituation phase of the experiment. A 2 X 2 mixed ANOVA (female vs. male, first two vs. last two habituation trials) revealed a main effect for trial block \(F(1, 14) = 193.36, p < .0001\) Mean\(_{\text{FIRST BLOCK}} = 11.41\) s, Mean\(_{\text{LAST BLOCK}} = 5.51\) s. There was no main effect for sex and no interaction. These results indicate that subjects did indeed habituate during the course of the experiment.

If the infants were able to learn the word-object associations and discriminate a change to a new CVC differing only in voicing, we would expect longer looking times to the “switch” trial than to the “same” trial. A 2 X 2 mixed ANOVA (female vs. male, same vs. switch) revealed no significant main effects or interactions (Mean\(_{\text{SWITCH}} = 6.41\) s, Mean\(_{\text{SAME}} = 7.06\) s). Thus, the infants familiarized to either the CVC [bɪn] or the CVC [pʰɪn] paired with the object “roundy”, and then tested with the contrasting word, failed to show a significant recovery in looking time. This indicates that they did not notice the switch to the new nonce word differing in voicing.

4.6 Discussion
The results of this experiment show that 14-month-olds fail to notice a switch in voicing when they are involved in learning a sound-meaning association. Furthermore, this result obtained when only a single pair was used in familiarization. One might hypothesize from this experiment, as well as the ones showing a lack of response to place, that restrictions on the complexity of lexical representations are causing the lack of response to the switch. If children initially represent minimally different words identically in the lexicon, then they would fail to notice the switch if this lexical representation is the one
tapped by these experiments. Speech perception or word discrimination tasks that do not require a link to meaning would tap more highly specified surface representations. However, another explanation would rely on processing. Being involved in a word learning task likely requires attentional resources that might limit the amount of attention paid to phonetic detail. In an attempt to tease apart these explanations, we designed an experiment that could falsify one version of the representational account.

5. Experiment 3: Voicing and Place Switch
One representational account of the results of Experiments 1 and 2 would be that 14-month-olds fail to notice place of articulation and voicing switches because they do not encode the relevant features in the lexical representations they construct for newly learned words. This account makes the prediction that a simultaneous switch of place and voicing should also fail to yield a response, since if neither place nor voicing is represented, then combining them should have no effect. Therefore, if it can be shown that 14-month-olds do succeed in responding to a simultaneous switch in place and voicing, then this representational account will be falsified. On the other hand, a processing explanation makes no such prediction. If processing load interferes with attention to phonetic detail, it could interfere with pick-up and use of both voicing and place features, predicting failure. Or, alternatively, because a 2-feature difference is likely more discriminable than a single feature difference, some versions of a processing account might predict that the greater dissimilarity created by switching both features would be more easily noticed.

Another motivation for running this experiment was to see if combining the potentially more salient contrast in place and voicing, with stimuli meeting English phonotactics, and the potentially simpler single pair methodology, would yield evidence of 14-month-olds using a consonantal contrast in word learning. If not, this provides further evidence of the robustness of the phenomenon identified by Stager and Werker (1997).

5.1 Participants
Sixteen 14-month-old infants completed this study, 8 girls and 8 boys (mean age, 14 months 18 days; range, 14 months 7 day to 14 months 27 days). All subjects were without apparent health problems and were exposed to English at least 90% of the time. An additional 18 infants were tested but were not included in the analyses because they were too fussy to continue (n = 7), their parents interfered in some way (n = 4), they were
reported to know *pin* \( (n = 2) \) or because of experimenter error. Subjects were recruited as in Experiments 1 and 2.

5.2 Stimuli

The stimuli were as in Experiment 1, except that the CVC syllables used were [dɪn] and [pʰɪn]. These syllables differ in both voicing and place of the initial consonant; (partially) voiced, unaspirated alveolar [d] vs. voiceless, aspirated bilabial [pʰ]. The nonsense label, [pʰɔk], was again presented during the pre- and post-test trials.

As in Experiment 2, only a single object, “roundy” was used during the habituation phase and test trials. The visual stimulus used during the pre- and post-test trials was again the “spinner”.

5.3 Equipment, Apparatus, and Procedure

The equipment and apparatus were identical to those used in Experiment 1. The procedure was identical to that in Experiment 2 with the exception that the parents were queried as to whether they used the words *pin* and *din* regularly.

5.4 Results

Reliability testing, performed as in experiment 1, yielded a Pearson product-moment correlation of equal to or greater than 95% for all subjects whose looking times were rescored, thus indicating that the measurement of looking time was reliable.

The test for fatigue over the course of the experiment (a 2 X 2 mixed ANOVA: female vs. male, pretest vs. posttest) yielded no main effects or an interaction, suggesting that the infants did maintain interest. The test for habituation (a 2 X 2 mixed ANOVA: female vs. male, first two vs. last two habituation trials)) did reveal a main effect for trial block \([F(1, 14) = 155.563, p < .0001\) Mean\(_{\text{FIRSTBLOCK}} = 13.08\) s, Mean\(_{\text{LASTBLOCK}} = 5.60\) s], with no main effect for sex and no interaction, indicating that subjects did indeed habituate during familiarization phase.

The analysis of the test trials (a 2 X 2 mixed ANOVA: female vs. male, same vs. switch) revealed no significant main effects or interactions (Mean\(_{\text{SWITCH}} = 6.91\) s, Mean\(_{\text{SAME}} = 6.76\) s). Thus, the infants familiarized to either [dɪn] or [pʰɪn] paired with the object “roundy”, and then tested with the contrasting word, failed to show a significant recovery in looking time, even though the words differed in both place of articulation and voicing.
5.5 Discussion
The results of this experiment show that when the switch involves both a place and a voicing distinction, 14-month-olds continue to fail to respond to it. This serves to further generalize the finding that subjects of this age fail to notice differences between minimally contrastive syllables when involved in a word learning task: as well as place and voicing contrasts on their own, a combined place and voicing switch is insufficient to evoke a response.

A finding that the subjects noticed the switch in this experiment would have falsified the representational account of these results outlined in the introduction to section 5. However, the lack of a response runs counter to neither the representational nor the processing account, so both do remain viable.

6. Partial integration of phonological contrasts
In the experiments reported in Stager and Werker (1997) and Werker et al. (2002), 14-month-old infants fail to respond to a switch between [b1] and [d1] when engaged in a word learning task. This result holds under a number of testing conditions. Stager and Werker (1997) used both a single sound-object pair, and two sound-object pairs in the habituation phase. Using two-sound object pairs, Werker et al. (2002) show that the switch is still not noticed when the objects are made more different from one another, and exposure time is increased (trials lengthened from 14 to 20 seconds, habituation criterion changed from 65% to 50%). Furthermore, in the experiments presented in this paper, neither adding a final nasal to the stimuli to bring them up to proper English word form (Experiment 1: [b1n] vs. [d1n]), nor using a voice switch (Experiment 2: [b1n] vs. [p1m]), or even a combined place and voice switch (Experiment 3: [d1n] and [p1n]), resulted in 14-month-olds showing a difference between switch and same trials. This held even with the potentially simpler single pair methodology, used in experiments 2 and 3.

These results suggest that the consonantal place of articulation and voicing contrasts are not yet fully integrated into the phonological system at 14 months, since they cannot be used to support learning a new meaning distinction. There is evidence that these contrasts have been at least partially acquired by English speaking infants at this age, insofar as the criterion for receptive phonological acquisition is the ability to link a sound and a meaning distinction. Swingley and Aslin (2002) used a visual fixation procedure to examine infants’ representation of well-known words (see also Swingley and Aslin 2000 for a similar study with older infants). In this procedure, two objects are shown on a computer screen. After a word is aurally presented that matches one of the objects, the subjects’ eye movements are monitored to determine how much time is spent
looking at the target. Swingley and Aslin (2002) compared accurate pronunciations of the target words to ones that differed minimally (e.g. [bejbi] vs. [vejbi] for ‘baby’). As with the older infants in Swingley and Aslin (2000), infants at a mean age of 15 months looked significantly longer at the accurate pronunciations. This study provides the first demonstration that infants at this age use consonantal place and voicing contrasts to encode meaningful words. While the differences between the accurate and “mispronounced” forms also included vocalic contrasts (e.g. [æp] vs. [æp] for ‘apple’), Swingley and Aslin (2002) report that the looking time difference held for all words, including those differing in consonantal place ([bɔl] vs. [gɔl] for ‘ball’ and [kʰri] vs. [pʰri] for ‘kitty’), and voicing ([dɔɡ] vs. [tʰɔɡ] for ‘dog’).

The infants tested by Swingley and Aslin (2002) are slightly older than those tested in the studies by Werker and colleagues (a mean of about 15 months as opposed to 14 months), and the contrasts they examined were not exactly the same as the labial/alveolar contrast used in the other studies. However, Fennell and Werker (2003) show that when 14-month-olds are habituated to pairings between [bɔl] and the object ‘ball’, and between [dɔl] and ‘doll’, they do notice a switch in sound-object pairing (these words share the same vowel in Canadian English, the dialect of the speaker of the stimuli, and of the environment of the infants). This provides an unequivocal demonstration that 14-month-olds do encode the same place distinction in well-known words that they ignore in learning new words.

There are a number of explanations one might give for the difference between these two sets of results. Here we follow Fennell and Werker (2003) and Werker and Curtin (in prep.) in ascribing the difference to processing load. The Swingley and Aslin (2000, 2002) and Fennell and Werker (2003) studies involve word recognition, in that well-known words were used, while those of Stager and Werker (1997), Werker et al. (2002), and the present paper, involve establishment of new sound-object pairings. It is plausible that word learning requires more cognitive resources, or attention, than word recognition, with the result that in such a task there are fewer resources (or less attention) available for phonological processing.

We suggest that the diminished availability of resources for phonological processing results in the construction of a phonological parse that is reduced in complexity (see Werker and Curtin 2004 for a somewhat different formalization of this idea). The complexity of a phonological structure is regulated by markedness constraints in Optimality Theory (OT; Prince and Smolensky 1993), as in earlier constraint-based theories of phonology. In OT, these constraints are universal, and whether or not a grammar enforces a constraint depends on its ranking with respect to conflicting
markedness and faithfulness constraints. Faithfulness constraints demand a match between structures at different levels of representation. For perception, we assume that faithfulness constraints apply between an initial phonetic parse of the acoustic signal, and the phonological parse (Boersma 1998, Pater 2004).\(^1\) To take the example of the labial/alveolar place contrast, the relevant markedness constraint would be one against labial consonants (*LABIAL; Prince and Smolensky 1993), while the faithfulness constraint would demand that place of articulation be identical between corresponding segments in the phonetic and phonological parse (IDENT[PLACE]; McCarthy and Prince 1999).

With *LABIAL ranked above IDENT[PLACE], the place contrast is neutralised, and both labials and coronals map to the unmarked coronal place. The illustrative tableaux in Table 2 show the choice of a phonological parse given the perceived strings [bI] and [dI]:

<table>
<thead>
<tr>
<th>Phonetic parse:</th>
<th>*LABIAL</th>
<th>IDENT[PLACE]</th>
</tr>
</thead>
<tbody>
<tr>
<td>[bI]</td>
<td>*!</td>
<td>*</td>
</tr>
<tr>
<td>[dI]</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

**Table 2: Tableaux for phonological parses with *LABIAL >> IDENT[PLACE]**

With this ranking, the contrast between [bI] and [dI] is merged at the phonological level. This ranking characterises the receptive phonological system of 14-month-olds, when under the cognitive load of word learning.

Boersma (1998) and Davidson et al. (2004) propose that in developing first and second language production respectively, the ranking of faithfulness constraints can vary depending on attention and other factors. A higher degree of attention leads to higher ranking of faithfulness constraints. This is an extension of proposals that treat variation as resulting from unranked constraints (see e.g. Anttila 1997, Boersma 1998, Boersma and Hayes 2001, as well as Demuth 1997, Boersma and Levet 2000, Pater and Werle 2001, and Curtin and Zuraw 2002 on acquisition). In the present context, we suggest that the ranking of IDENT[PLACE] and *LABIAL varies depending on the resources available for

\(^1\) The proposals in Boersma (1998) and Pater (2004) differ on what markedness constraints affect phonological perception. In Pater (2004), there is a single set of markedness constraints, with perception- and production-specific faithfulness. In Boersma (1998), there is a richer set of markedness constraints, with ones specific to each of perception, production, and recognition. However, though this is not noted in Pater (2004), they share the idea that faithfulness applies between the perceived acoustic string and the receptive phonological parse.
phonological processing. With fewer resources available, the ranking in shown in Table 2 obtains, and the structures are simplified. With more resources available, faithfulness is promoted relative to markedness, and the contrast is maintained phonologically, as illustrated in Table 3. This ranking characterises the receptive phonological system of infants that are not under the cognitive load of word learning, or for whom word learning has become less cognitively demanding.

Partial integration of a contrast, then, is here conceived of as a variable ranking between markedness and faithfulness constraints, which is affected by cognitive load. When the contrast becomes fully integrated, the ranking of the faithfulness constraint is fixed above the markedness constraint, and the marked structure is maintained in the phonological parse, regardless of processing load. Under this view, acquisition of a contrast is not an all or nothing affair, but proceeds gradually.

This account makes two predictions about the interaction of task complexity and phonological complexity that can be investigated in future research. First, an increase in task complexity should result in an increase in the effects of phonological markedness constraints. While this prediction remains to be systematically investigated, there are several results that seem to support it. Werker et al. (2002) show that 17-month-olds are able to use a place distinction to distinguish newly learned words, and similar results have been obtained using different testing procedures (Bailey and Plunkett, 2003). These results contrast with Barton’s (1976) finding that newly learned words were not well discriminated in a minimal pair task, and with the general finding in such studies that contrasts emerge gradually over the course of development. Since the minimal pair studies do have higher task demands than the ‘Switch’ task, this may be taken as evidence of the greater effects of markedness constraints under higher cognitive load. More conclusive evidence could be obtained on this score by extending the looking time studies to more “difficult” contrasts in tests with older infants, and by conducting more carefully controlled minimal pair-style tasks. The other body of evidence that seems
congruent with this prediction comes from studies of adult cross-language speech perception, which show greater influence of the native language under increased task demands (e.g. Werker and Tees 1984b, Dupoux et al. 2002). Showing that the use of lexical phonological contrasts in cross-language perception is also influenced by task demands would provide strong support for a partial integration view.

Second, it is predicted that when task difficulty is held constant, more marked structures should cause greater difficulty in discrimination. Again, this seems to be evidenced by the classical minimal pair tasks, but these typically suffer from methodological flaws that make it difficult to interpret the results (see e.g. Barton 1976, 1980). Here again, extension of the looking time experiments to more marked structures would be particularly revealing. The contrast that has been most studied thus far, [b] vs. [d] in word-initial onset position, is one of the least marked of the consonantal contrasts (Jakobson 1941/68: 73). Some examples of more marked contrasts would be instances of the same place contrast in other contexts, such as in word-final coda position, or in fricatives ([f] vs. [s]). Cross-linguistically, and in the development of production, the presence of either of these more marked contrasts implies the presence of the onset stop place contrast. These implicational universals are captured in Optimality Theory either by having markedness constraints specific to the more marked contrasts (e.g. *CODALABIAL), or faithfulness constraints specific to the unmarked one (e.g. IDENT-PLACE-ONSET). With such constraints, it is impossible to construct grammars that allow a place contrast only in the marked position. Assuming that the same types of constraints apply in receptive parsing (see Pater 2004), then we would predict that discriminating a contrast in a marked context should imply the ability to discriminate it in an unmarked context.

The finding of differences between relatively marked and unmarked contrasts in word learning tasks would support a model such as the present one, in which phonological constraints mediate between the perceived acoustic string and the lexicon. However, it would be possible to account for many of the same results in terms of acoustic salience (Fennell and Werker 2003; Werker and Curtin 2004), given that unmarked contrasts tend to be acoustically salient. One way of disambiguating these accounts would be to follow the logic of Experiment 3. The phonological account presented in this section, like the level-based account, would not predict additive effects of contrasts, so if they are found, this would support the account based on salience. However, the possibility of null effects, like the one found in Experiment 3, makes this a difficult research strategy to pursue. Finding conclusive evidence that phonological markedness is necessarily at issue also poses a significant methodological challenge, though recent work on adult perception suggests ways to meet it (Moreton 2002).
The idea that the ranking of grammatical constraints can be affected by task demands runs contrary to the traditional competence/performance distinction. Indeed, in syntactic acquisition studies, the general tack in assessing the state of a learner’s developing competence is to minimize task demands, so as abstract from performance factors. Applying this approach here might lead one to conclude that the place and voice distinctions are acquired phonologically by 14-15 months, given the results of Swingley and Aslin (2000, 2002) and Fennell and Werker (2003). One difficulty with this view is that it is at odds with how phonological perception is typically conceived of in the literature on adult processing: as relatively impervious to task demands, in comparison with acoustic or phonetic perception. Clearly, these segmental contrasts are not robustly maintained at 14-15 months. An especially unattractive aspect of this way of maintaining a strict competence/performance distinction is that it relegates phonology to an extremely limited role in accounting for the course of acquisition. We favor the present approach because it yields a set of testable predictions across a variety of domains.

7. Conclusions
We began this paper by contrasting two diametrically opposed positions on the perceptual acquisition of phonological contrasts. The first is that they are acquired gradually, in much the manner that they are acquired in production. The second is that there are no constraints on the phonological complexity of early perceptual representations, beyond those imposed by the ambient language. In studies using looking time as a measure of sound-meaning pairing, we find evidence that supports the first of these positions, in that fourteen-month-olds do not respond to a switch in consonantal place or voicing when engaged in a word learning task. However, we have also reviewed evidence that seems more in line with the second, in that when well-known words are used in such studies, even 14-to-15-month-olds respond to place and voice switches. In addition, we have noted that insofar as the ability to represent contrasts in memory is indicative of phonological acquisition, then infants as young as 71/2-months can be said to have acquired a place of articulation contrast. The interpretation of these findings offered here is that before being fully acquired, contrasts are partially integrated into the phonological system, during which time their maintenance is affected by processing demands, such as the establishment of sound and meaning pairings.
References


Fennell, Chris and Janet Werker. 2003. Early word learners’ ability to access phonetic detail in well-known words. Language and Speech 46(2-3). 245-264.


