CAMBRIDGE
UNIVERSITY PRESS

BRIEF RESEARCH REPORT

Fourteen-month-olds' sensitivity to acoustic salience in minimal pair word learning

Stephanie L. ARCHER[1]* and Suzanne CURTIN[2]

[1]Department of Linguistics, University of Alberta and [2]Department of Psychology, University of Calgary
*Corresponding author. E-mail: slarcher@ualberta.ca

Abstract

During the first two years of life, infants concurrently refine native-language speech categories and word learning skills. However, in the Switch Task, 14-month-olds do not detect minimal contrasts in a novel object–word pairing (Stager & Werker, 1997). We investigate whether presenting infants with acoustically salient contrasts (liquids) facilitates success in the Switch Task. The first two experiments demonstrate that acoustic differences boost infants' detection of contrasts. However, infants cannot detect the contrast when the segments are digitally shortened. Thus, not all minimal contrasts are equally difficult, and the acoustic properties of a contrast matter in word learning.

Keywords: First Language Acquisition; Word Learning; Speech Perception

Introduction

Infants' ability to discriminate sound contrasts develops over the first two years of life, beginning with a fairly unrestrained ability to perceive differences in both native and non-native contrasts (Eimas, Siqueland, Jusczyk, & Vigorito, 1971) to perceiving contrasts in their native language (Werker & Tees, 1984). As this perceptual honing of native vowels around 6 months (Polka & Werker, 1994) and native consonants around 10 months (Werker & Tees, 1984) becomes more attuned, infants begin recognizing highly familiar words (Bergleson & Swingley, 2012), and by 12 months they are mapping novel labels to objects that conform to their native language phonology (Curtin, 2011; MacKenzie, Curtin, & Graham, 2012). However, discrimination of speech sound contrasts is dependent not just on experience, but also on factors such as the position in the word of the contrast (Archer, Zamuner, Engel, Fais, & Curtin, 2016), and the salience of the contrast (Narayan, Werker, & Beddor, 2010). Thus, while infants are increasing their knowledge of the sound system, and becoming better at discriminating native contrasts (Kuhl, Stevens, Hayashi, Deguchi, Kiritani, & Iverson, 2006) in word beginnings, they still struggle with those contrasts that are not as perceptually robust, and those in less salient positions within a word. This pattern is also observed in the second year of life when infants are learning new word–object mappings.

The ability to form word–object associations by 12 months is limited by how distinctive the labels are from one another. For example, when presented with novel

minimally contrastive word–object pairs (e.g., *bih* [bɪ] and *dih* [dɪ] or *bin* [bɪn] and *din* [dɪn] in Pater, Stager, & Werker, 2004), 14-month-olds are unable to access place information. Results are similar with voice (*pin* [pʰɪn]–*bin*), and voice–place contrasts (*pin–din*) (Pater *et al.*, 2004). These findings are perplexing, given younger infants' ability to discriminate these sound contrasts and learn words (Werker & Curtin, 2005). It has been proposed that when presented with minimally contrastive novel labels, infants have limited access to available resources (Stager & Werker, 1997). This problem is overcome by simply accruing experience with language (around 17 months; Werker, Fennell, Corcoran, & Stager, 2002), but younger infants require additional information. For example, 14-month-olds can access minimal contrasts when provided with referential information (Fennell & Waxman, 2010) or distributional information (Thiessen, 2007), or when task demands are eased (Yoshida, Fennell, Werker, & Swingley, 2009). To further understand challenges in word learning, we explore whether some minimal contrasts are salient enough due to their acoustic properties to be detected. Specifically, we ask whether acoustic properties of a sonorant contrast can ease the task of mapping minimally contrastive forms to novel objects.

In a novel word–object association task (i.e., Switch Task), novice word learners must identify the contrast between pairs, map the forms to objects, and retrieve this information. If a minimal contrast is acoustically salient, perhaps the contrast itself is enough to boost infants' success in this task. Acoustic salience, in this context, can be defined as the acoustic factors that might contribute to increased perceptibility of the listener. Support for this comes from a handful of studies demonstrating that the acoustic properties of a minimal contrast influences whether infants can access the subtle phonetic detail. Curtin, Fennell, and Escudero (2009) investigated whether infants would detect minimal vowel contrasts because vowels are acoustically salient sound segments containing their own formant structure. To test this, one group of 15-month-olds was habituated to word–object pairs using the minimal pairs *deet* [dit] and *doot* [dut] and another group to *deet* [dit] and *dit* [dɪt]. Infants exposed to the *deet–dit* contrast succeeded at detecting the mismatch. However, the infants exposed to *deet–doot* did not. These results were attributed to the acoustic differences between the vowels; 15-month olds respond to vowel height contrasts (a difference in F1) but not vowel backness contrasts (a difference in F2), suggesting that an F1 contrast is more salient to infants than a difference in F2 (Curtin *et al.*, 2009). Indeed, there is more acoustic energy in F1 (Lacerda, 1993, 1994).

The position in which segments surface can also enhance the salience of a contrast. For example, when the minimal contrastive information is contained in a stressed medial syllable of a trisyllabic nonword (i.e., *lebóna* [lebóna], *ledóna* [ledóna]) 14-month-olds succeed at detecting the switch (Archer, Ference, & Curtin, 2014). This is not the case if the ɪɴɪᴛɪᴀʟ stressed syllable contains the contrast, as in *bólena* [bólena] and *dólena* [dólena]. Archer *et al.* (2014) argue this is due to a difference in the locus of the second formant transition (F2), the primary source of place information for stops (Delattre, Liberman, & Cooper, 1955) where a larger difference is observed in the medial condition.

The acoustic properties of contrasts can impact infants' ability to access this information when presented, at test, with a mismatched object–label pairing, suggesting that the task demands are reduced when the contrast is acoustically salient (Curtin *et al.*, 2009; Archer *et al.*, 2014). Though these studies demonstrate that acoustic enhancement increases infants' ability to access a minimal contrast, it has yet to be demonstrated with consonants in word-initial position. This is surprising

because word-initial position is arguably the most salient position (Wang & Seidl, 2015). It might be the case that word-initial position is salient in some tasks, such as discrimination (Zamuner, 2006), but is outweighed by resource limitations when the task is more challenging (see Archer *et al.*, 2014). Here, we explore salient minimal pairs using sonorant consonants in word onset position to determine whether a place of articulation contrast, based on differences in formant signatures, is more easily detected in a word learning task than the contrasts tested in previous studies.

The information that determines the place of articulation in stops, like /b/ or /d/, is housed within the first few milliseconds of the following vowel (FORMANT TRANSITION). Shown in Figure 1a, bilabial stops, (/b/) are defined by the second and third formants (F2 and F3) rising into the vowel /ɪ/, while the locus of the formant transition in the coronal (/d/) is higher in frequency (Hz) and therefore remains unchanged as it enters the vowel. English liquids (/l/ and /r/) are sonorant, and rather than relying on a formant transition into the vowel as with stops, liquids have
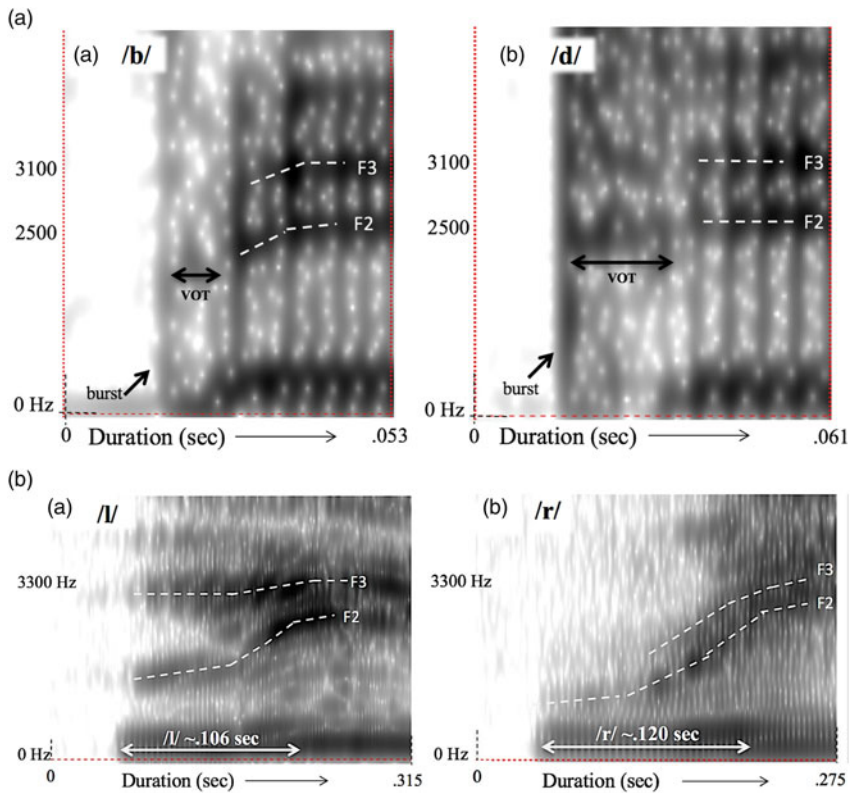

Figure 1. (a) Example: Spectrogram (a): .053 sec window includes /b/ to beginning of vowel /ɪ/ (*bin*). Spectrogram (b): .061 sec window includes /d/ to beginning of vowel /ɪ/ (*din*). These examples illustrate the components of voiced stops in which place information (F2 and F3) is very short in duration, as compared to those in Figure 1b. Formants and formant transitions are highlighted. (b) Example: Spectrogram (a): .315 sec window includes /l/ to beginning of vowel /i/ (*leet*). Spectrogram (b): .225 sec window includes /r/ to beginning of vowel /i/ (*reet*). These examples illustrate the duration of the segment and place information (F2 and F3), as compared to those in Figure 1a. Arrows demonstrate duration of consonants /l/ and /r/. Formants and formant transitions are highlighted.

their own formant structures. That is, while F2 and F3 are still the primary indicators of each liquid, the formants which differentiate /l/ and /r/ are contained within the segment and longer in duration, thus providing more time for the changing structure to be detected. In this series of experiments, we explore whether infants detect a perceptually salient contrast in word-initial position (Experiment 1), and when it is the second member of an onset cluster (Experiment 2). Finally, we examine whether the duration of the formant pattern is important for distinguishing the contrast (Experiment 3). We predicted that the contrast between sonorant segments (i.e., liquids) will be perceptually salient to 14-month-olds and that they will therefore detect the mismatches in a newly presented word–object pairing.

Experiment 1

Fourteen-month-olds were habituated to two novel word–object pairings in which the minimal contrast between the words was in the initial consonant. We predicted that infants would detect the mismatch in the Switch trial based on the acoustically salient contrast /l/ and /r/ (i.e., *leet* and *reet*).

Methods
Participants

Sixteen 14-month-old infants (*M*: 14.51, *SD*: 0.321; Range: 14.00–14.98) from monolingual English homes participated in this study. An additional 15 infants were tested but not included in the analysis (see 'Appendix A' for a detailed report). Our sample size was determined by using a stopping rule (16 infants who succeeded at completing the task) based on prior studies using the Switch Task (e.g., Fennell & Waxman, 2010; MacKenzie *et al.*, 2012).

Stimuli

A female native speaker of English recorded the auditory stimuli consisting of two one-syllable words, *leet* [lit] and *reet* [rit], using infant-directed speech. Acoustic measurements of duration are presented in Table 1. Habituation and test trials contained 16 tokens of either *leet* or *reet* made up of 4 unique tokens semi-randomized within blocks. Each trial was 22 seconds long with an interstimulus interval (ISI) of 750 ms. Each of the novel words was associated with a novel object (Stager & Werker, 1997).

Procedure

We used a modified version of the Switch procedure (Werker, Cohen, Lloyd, Casasola, & Stager, 1998). Fourteen-month-olds were habituated to two word–object pairs (see Figure 2) and then tested on their ability to detect a mismatch in the pairing. A PRE-TEST trial began the experiment with a non-related novel word (*wug*) through speakers paired with an object on a screen. Subsequently, each trial began with a silent attention-getter used to re-engage infants between trials. Once engaged, the HABITUATION PHASE started with a novel object appearing with an accompanying novel word: *leet* paired with Object A and *reet* paired with Object B, counterbalanced across groups of infants. Four trials of word–object pairs comprised a 'block' in a semi-random order (e.g., ABAB, BAAB) and eight semi-randomized blocks

Table 1 Acoustic measurements of stimuli

	Experiment 1		Experiment 2		Experiment 3	
	leet	*reet*	*bleet*	*breet*	short *leet*	short *reet*
Word (sec)	.625 (.017)	.621 (.031)	.504 (.023)	.526 (.030)	.488 (.022)	.486 (.029)
Liquid + formant transition (sec)	.106 (.025)	.120 (.043)	.076 (.013)	.097 (.024)	.071 (.012)	.099 (.015)
F2 onset liquid (Hz)	2525.2 (202.6)	1525.7 (84.8)	1714.7 (142.8)	1390.8 (65.1)		
F2 offset liquid (Hz)	2752.3 (347.1)	2207.1 (319.2)	2332.2 (60.1)	2125.9 (255.3)		
F3 onset liquid (Hz)	3266.0 (174.5)	2084.1 (126.5)	2819.0 (169.2)	2181.8 (270.8)		
F3 offset liquid (Hz)	3329.3 (80.5)	2811.2 (468.1)	3075.2 (267.9)	2813.4 (288.0)		

Note. Parentheses denote standard deviation (SD). Stimuli for Experiment 3 was created by splicing out the /b/ onset in Experiment 2, therefore we have similar means for F2 and F3 between the two experiments.
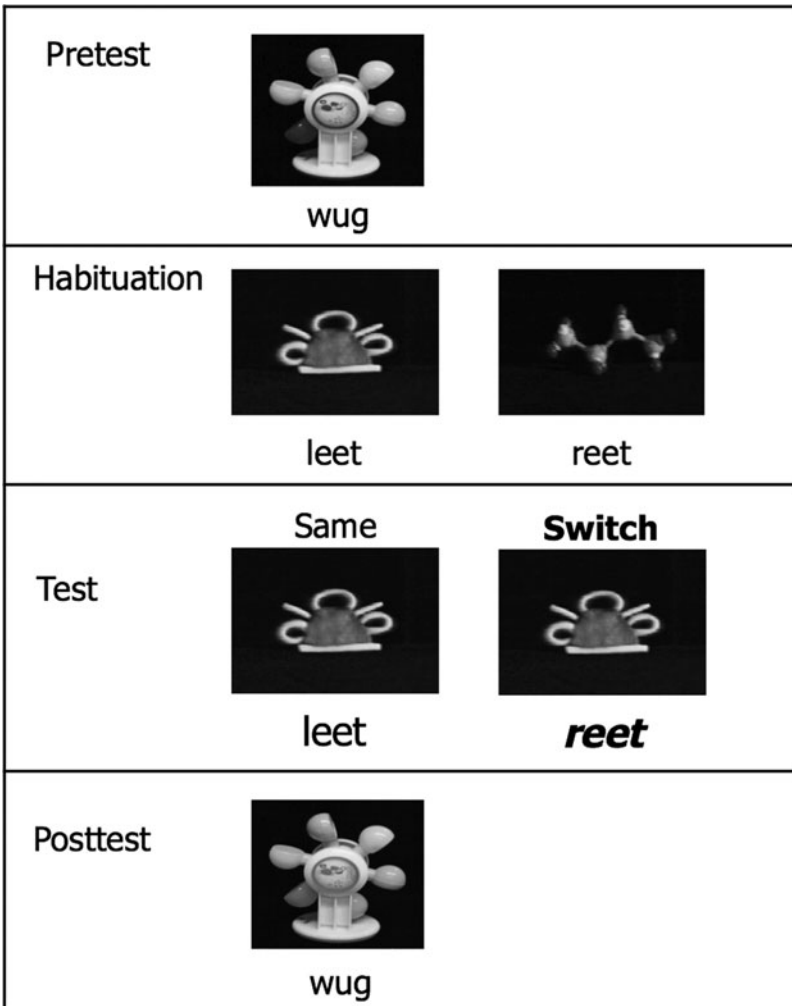
Figure 2. Modified version of the Switch procedure (Werker *et al.*, 1998). This figure illustrates the progression of this procedure in four phases. The PRE-TEST engages the infant using a novel word and object unrelated to the body of the experiment. The HABITUATION PHASE consists of 8 to 24 trials of novel word–object pairs (e.g., *leet* – Object A, *reet* – Object B, or vice versa). Two trials comprise the TEST PHASE: Same and Switch. The Same trial is identical to one of the two habituation trials. The Switch trial is a mismatch in which, for example, if *leet* was paired with Object A during the habituation phase, a Switch trial might pair *reet* with Object B. The POST-TEST ends the experiment.

comprised a habituation order, across participants. All trials were initially coded online to determine habituation. During the habituation phase, the average looking time for a particular infant across the first four-trial block was compared to subsequent four-trial blocks until looking time decreased to a pre-set criterion of 65% (as in Stager & Werker, 1997; Fennell & Werker, 2003). Infants were exposed to a minimum of eight and a maximum of 24 habituation trials.

The TEST PHASE began immediately after habituation and included two test trials: Same and Switch. Same trials were identical to pairs in the habituation phase and Switch trials

included a mismatch between object and label (see Figure 2). During the habituation phase, infants lose interest in the repetition of both word–object pairs. If mismatched stimulus is detected, infants will likely recover interest during the Switch trial. That is, the Switch trial is different from all other trials so it should be surprising to infants and therefore we expect longer looking times compared to the Same trial. Order of test trials and specific word–object pairing were counterbalanced across infants. After the test phase, a post-test trial, identical to pre-test, ended the procedure.

Apparatus

Testing took place in a 2.74 × 1.82 m quiet, dimly lit room. Infants sat on their parents' laps facing a screen (122 cm wide, 91.5 cm high) approximately 1.5 m away. Images were projected onto the screen via a NEC LT245 projector. Auditory stimuli were delivered at 65 ± 5 dB over a Bose 101 speaker located directly below the screen. Infants were recorded using a Sony DCRDVD92 digital video camera.

Parents wore Bose True Noise-Cancelling Headphones, through which music was played. Habit X 1.0 (Cohen, Atkinson, & Chaput, 2004), run on a Macintosh Power PC G5, was used to order digitized audio and visual stimuli presentations and collect looking time data. The experimenter, who was blind to the auditory stimuli and type of trial, monitored the infants' looking behaviour via a closed-circuit television system from an adjacent testing room and coded online by pressing a designated key. Video recordings were coded offline using Supercoder (Hollich, 2005) and used as data in the analysis. Cronbach's alpha was calculated as the reliability statistic. To measure inter-rater reliability, 25% of the data ($n = 4$ infants) were coded by a second coder. Data reliability required equal or above .97, and this level of agreement was achieved (Cronbach's $\alpha = .999$).

Results

We first determined whether infants recovered to the post-test by comparing looking times to the last habituation block (M: 7.40, SD: 2.62) with looking times to the post-test (M: 19.17, SD: 3.58) There was a significant effect of trial ($t(15) = -11.685$, $p < .001$, $d = 2.921$), suggesting that infants maintained interest in the task and recovered to the post-test.

Our primary comparison was infants' looking times during the Same and Switch test trials. A pairwise t-test showed a significant difference between the Same trial (M: 6.83, SD: 3.95) and Switch trial (M: 8.83, SD: 4.67) ($t(15) = -2.253$, $p = .040$, $d = 0.594$), demonstrating that infants were successful at detecting the mismatch at test (see Figure 3). Of the 16 participants, 12 looked longer at the Switch trial.

Discussion

The results of Experiment 1 suggest an acoustically salient contrast can override the challenge faced by novice word learners when presented with minimally different word forms (Stager & Werker, 1997). Although 14-month-olds have difficulty detecting some minimal pairs (e.g., *bin–din*; Pater *et al.*, 2004), if the contrast is salient then infants use the acoustic information to detect a mismatch without requiring the aid of referential (e.g., Fennell & Waxman, 2010) or distributional (e.g., Thiessen, 2007; Thiessen & Yee, 2010) information. Thus, the nature of the contrast alone can boost infants' ability to detect it in a Switch Task. However, it is not known what specific acoustic properties of the contrast are driving infants' ability to distinguish the minimal
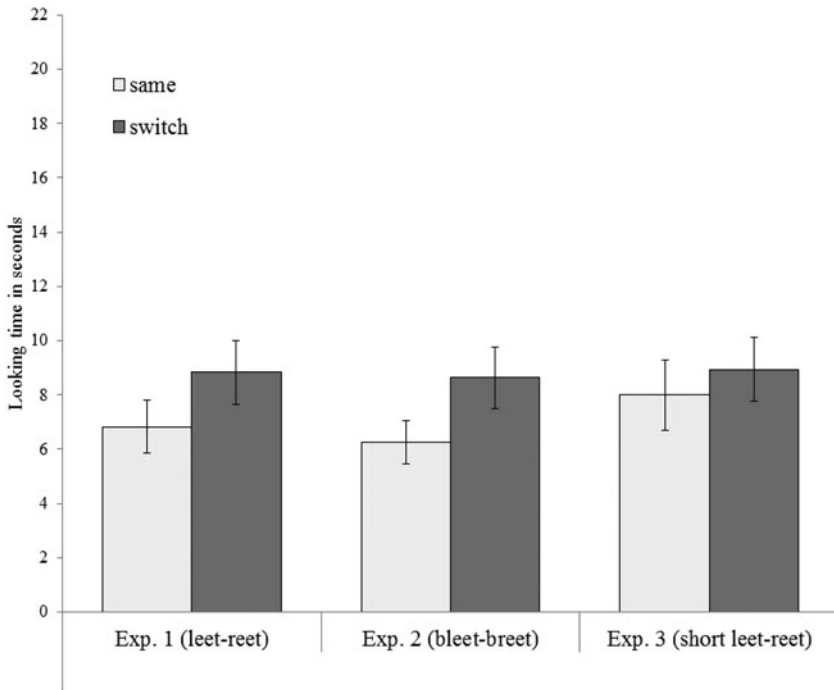
Figure 3. Mean looking time in seconds for Experiment 1 (*leet–reet*), Experiment 2 (*bleet–breet*), and Experiment 3 (shortened *leet–reet*) contrasts. Light grey denotes looks to Same trial (SAME) and dark grey denotes looks to Switch trial (SWITCH). Error bars denote SE.

pair *leet–reet*. That is, infants could be relying on the overall duration of the segment, and/or simply having a formant pattern is sufficient for success in this task. In Experiment 2, we investigate whether infants still access the contrastive information in a minimal pair when the liquid contrast is embedded in a complex onset. When a liquid is part of a complex onset, the duration is naturally shortened by the first element of the onset (i.e., /b/).

Experiment 2

We tested 14-month-olds using stop–liquid onsets (i.e., *bleet–breet*), in which the contrast is the second element embedded within a complex onset. As a result, some characteristics of the liquids change (see Figure 4 and Table 1 for acoustic measurements of stimuli). In this experiment, we specifically focus on the duration of the liquid contrast to determine whether it is a factor when infants are identifying contrasts. In this experiment, if infants do not detect the switch, then the acoustic or positional differences are no longer salient enough to outweigh resource limitations.

Methods

Participants

Sixteen 14-month-old infants (*M*: 14.45, *SD*: 0.263; Range: 14.10–14.92) from monolingual English homes participated in this study. Twelve additional infants were
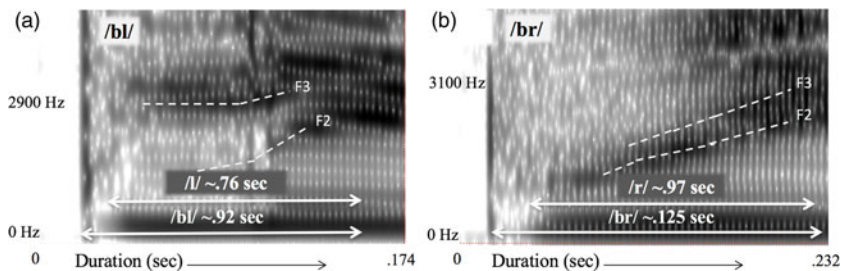
Figure 4. Example: Spectrogram a: .174 sec window includes /bl/ to beginning of vowel /i/ (*bleet*). Spectrogram b: .232 sec window includes /bl/ to beginning of vowel /i/ (*breet*). These examples illustrate the duration of the liquid contrast to compare to those in Figure 1b.

tested and are reported in 'Appendix A'). The same stopping rule applied as in Experiment 1.

Stimuli

The same speaker as in Experiment 1 recorded stimuli in infant-directed speech: *bleet* [blit] and *breet* [brit]. Each trial was 21 seconds long with an ISI of 750 ms. Onset measurements are reported in Table 1.

Procedure and apparatus

The procedure and apparatus for this experiment are identical to Experiment 1.

Results

As in Experiment 1, we ran an analysis of infants' recovery to the post-test (*M*: 17.51, *SD*: 4.30) after the last block of habituation trials (*M*: 8.72, *SD*: 2.28). There was a significant effect of trial ($t(15) = -7.499$, $p < .001$, $d = 1.875$), indicating that the infants recovered to the post-test.

As in the previous experiment, a pairwise *t*-test comparison of Same and Switch trials showed a significant difference between looks to the Same trial (*M*: 6.25, *SD*: 3.16) and Switch trial (*M*: 8.63, *SD*: 4.56) ($t(15) = -3.146$, $p = .007$, $d = 0.787$) (Figure 3). Thirteen of the 16 infants looked longer to the Switch. Inter-rater reliability parameters were identical to those in Experiment 1 and yielded a score exceeding .97 (Cronbach's $\alpha = .999$).

Discussion

Fourteen-month olds detect the /l/–/r/ contrast, even in the second position of an complex onset. Though the liquids in these onsets were shorter in duration (compared to the *leet–reet* contrast), infants nonetheless detected the mismatch in the Switch trial, suggesting that consonants with formants are salient regardless of duration. Alternatively, because the first element of the cluster includes anticipatory coarticulation information about the upcoming segment, infants may be relying on this additional source of information to distinguish between the liquids. To tease this apart, we tested whether infants detect the mismatch when the liquid maintains the same duration as when it is embedded in an onset, but without coarticulatory information in the preceding stop.

Experiment 3

To test whether infants were attending to the duration of the formant structure of the segments (Experiment 1) or were aided by coarticulatory information contained within the cluster (Experiment 2), we created a singleton onset with the duration of a liquid in a complex onset (i.e., *bleet–breet*) without the coarticulatory information by removing the /b/; see Table 1. We predicted that in the absence of coarticulatory cues, infants will still have available sufficient information in the formants, even though they are shorter, to distinguish between the contrasts in a word learning task. That is, we predicted infants will detect a switch in the word–object pairing.

Methods

Participants

Sixteen 14-month-old infants (*M*: 14.57, *SD*: 0.273; Range: 14.00–14.98) from monolingual English homes participated in this study. A further eight were not included in the analysis (see 'Appendix A' for detailed report).

Stimuli

The stimuli were created by splicing out the stop from the *bleet* and *breet* tokens using Praat (Boersma & Weenink, 2010). This method of altering stimuli preserved the duration of the liquid within the onset while creating a singleton onset (see Table 1 for measurements). Three adult speakers of English reviewed each token and reported that they perceived a singleton liquid. Each trial was 20 seconds long with an ISI of 750 ms.

Procedure and apparatus

The procedure and apparatus were identical to Experiments 1 and 2.

Results

Following Experiments 1 and 2, we ran an analysis confirming infants' recovery from the last four habituation trials (*M*: 8.66, *SD*: 2.85) to the post-test (*M*: 17.08, *SD*: 4.20). We again found a significant effect between trials ($t(15) = -8.267$, $p < .001$, $d = 2.067$).

We used a pairwise *t*-test comparison of infants' looking times during the Same and Switch trials as our primary analysis. This analysis showed no significant difference between the Same (*M*: 8.00, *SD*: 5.22) and Switch trials (*M*: 8.93, *SD*: 4.70) ($t(15) = -0.785$, $p = .444$) (Figure 3). A Cronbach's $\alpha = .987$ score exceeded the reliability standard .97.

Contrary to our prediction, infants of 14 months were unable to detect the contrast during the test phase. When the initial stop information was removed from the words *bleet* and *breet*, infants were presented with a shortened onset liquid, contrasting only in duration with the natural stimuli in Experiment 1 (*leet–reet*). These results suggest that infants struggled with this contrast, with only six of the 16 infants looking longer during the Switch trial.

To determine whether there were any significant interactions across experiments, we compared Experiment 1 (*leet–reet*) to Experiment 3 to see if duration alone was driving infants' success in Experiment 1, and Experiment 2 (clusters) to Experiment 3 to see if coarticulatory cues were driving success in Experiment 2. There were no significant interactions (all *p*s >.05; please see 'Appendix B'), suggesting that we cannot

conclusively determine whether duration or coarticulatory cues are driving performance. Rather we can only conclude that in the absence of additional acoustic information, infants are faced with a greater challenge in this task.

General Discussion

In this series of experiments, we found that 14-month-old infants succeed at learning new pairings between novel objects and words when the labels contain a minimal acoustically salient contrast. Specifically, the formant structures and coarticulatory cues provide the necessary salience for infants to detect minimal pairs. Importantly, when these cues are weakened, as was the case in our third experiment, infants no longer detect a mismatch in the pairing. However, exactly why infants do not detect the switch in Experiment 3 requires further exploration, since it could be because of the lack of sufficient duration or coarticulatory cues, or even the artificially created context. Recall that, by 10 months, infants discriminate only those contrasts experienced in their native language (Werker & Tees, 1984). In Experiment 3, the artificially created stimuli could be treated as non-native, thereby constraining infants' willingness to map these forms to objects (Mackenzie, Graham, Curtin, & Archer, 2014; May & Werker, 2014). Perhaps if infants have experience with a shortened liquid segment at word onset, they might succeed, or if the infants were provided with additional support through referential information (Mackenzie *et al.*, 2014) they might accept these forms as possible labels.

This finding complements other studies (Curtin *et al.*, 2009; Archer *et al.*, 2014) that demonstrate infants' ability to access fine phonetic detail in minimal pairs when the particular contrast is salient due to various acoustic factors. Though previous onset contrasts (e.g., *bin–din*) were inaccessible to 14-month-olds, the results of our experiments show that this difficulty is not related solely to position or minimal pairs in general. Here we found that, given minimal pairs that contrast in singleton liquids (*leet–reet*) or liquids in complex onsets (*bleet–breet*), infants are able to access the contrasting information, illustrating that acoustic salience is indeed a factor in word learning.

Our interpretation of these results supports the notion that not all contrasts are equally salient or detectable, and that certain tasks at certain stages of phonological development can impede infants' ability to access different types of information (see Werker & Curtin, 2005; Curtin, Byers-Heinlein, & Werker, 2011). These findings are the first demonstration of minimal contrasts in onset position that overcome the resource limitations in 14-month-olds based only on the acoustic properties of the specific contrasts (in this case, *leet–reet* (duration) and *bleet–breet* (coarticulation)) in the Switch task (Werker *et al.*, 1998). This suggests that some cues may provide extra processing time enhancing infants' access to contrastive information. Related to this, Archer *et al.* (2016) showed that discrimination of stop contrasts in coda position were perceived earlier if the acoustic properties of the contrasts were salient (e.g., /ab/–/ag/ but not /ap/–/ak/). Together, these findings suggest that success in a task is largely due to acoustic properties of the contrast.

Contrasts that are acoustically and perceptually salient are easier for novice word learners to access. In an exemplar-like system, tokens of linguistic input (specifically, word forms) are stored in terms of their phonetic similarity (Werker & Curtin, 2005; Curtin *et al.*, 2011). In illustration, tokens such as *fep* and *wug* contrast on a number of dimensions, so clusters of *fep* tokens and *wug* tokens can be stored separately.

This eases access to these word forms at test (MacKenzie, Graham, & Curtin, 2011). However, in the case of minimal pairs, infants are challenged with accessing forms that differ by a single change in segments. In terms of storage, similar-sounding word forms are stored close together or even in potentially overlapping exemplar space. Access in these cases becomes difficult because, in a word–object association task (i.e., Switch) other sources of information are not included in the experience. In support of this notion, 17-month-olds successfully learn minimal pairs when habituated first to low-density neighbourhoods (Storkel & Rogers, 2000; Hollich, Jusczyk, & Luce, 2002). When phonological neighbourhoods have a high density, access becomes more difficult because of competition between closely stored tokens. In the case of our experiments, acoustically salient contrasts are more likely to be distinct from one another, thus separating word form clusters in exemplar space, allowing easier access.

Findings from studies show that acoustic salience (specifically duration and/or coarticulatory information) impacts minimal pair word learning, along with context, language experience, and developmental stage. By investigating the various factors that influence word learning in the early stages of language development we come further towards understanding the interplay between the acoustic signal, the resources at hand, and the demands of the task. That is, multiple factors continually influence infants' mapping and retrieval of information while, dynamically, the system incorporates new information which feeds back into itself, creating stronger representations and a growing vocabulary, suggesting that studies need to take into account numerous factors, including the acoustic properties of the sound segments, in order to predict and provide a comprehensive view of early word learning and phonological development (Werker & Curtin, 2005).

References

Archer, S. L., Ference, J. D., & Curtin, S. (2014). 14-month olds succeed at learning minimal pairs in stressed syllables. *Journal of Cognition and Development*, *15*(1), 110–22.

Archer, S. L., Zamuner, T., Engel, K., Fais, L., & Curtin, S. (2016). Infants' discrimination of consonants: interplay between word position and acoustic saliency. *Language Learning and Development*, *12*(1), 60–78.

Boersma, P., & Weenink, D. (2010). Praat: a system for doing phonetics by computer (Version 5.0.29) [Computer Software]. Retrieved from http://www.praat.org.

Cohen, L. B., Atkinson, D. J., & Chaput, H. H. (2004). Habit X: a new program for obtaining and organizing data in infant perception and cognition studies (Version 1.0). Austin: University of Texas.

Curtin, S. (2011). Do newly formed word representations encode non-criterial information? *Journal of Child Language*, *38*(4), 904–17.

Curtin, S., Byers-Heinlein, K., & Werker, J. F. (2011). Bilingual beginnings as a lens for theory development: PRIMIR in focus. *Journal of Phonetics*, *39*, 492–504.

Curtin, S., Fennell, C. T., & Escudero, P. (2009). Weighting of acoustic cues explains patterns of word–object associative learning. *Developmental Science*, *12*, 725–31.

Delattre, P. C., Liberman, A. M., & Cooper, F. S. (1955). Acoustic loci and transitional cues for consonants. *Journal of the Acoustical Society of America*, *27*, 769–73.

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, *171*(968), 303–6.

Fennell, C. T., & Waxman, S. (2010). What paradox? Referential cues allow for infant use of phonetic detail in word learning. *Child Development*, *81*(5), 1376–83.

Fennell, C. T., & Werker, J. F. (2003). Early word learners' ability to access phonetic detail in well-known words. *Language and Speech*, *46*, 245–64.

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

Hollich, G., Jusczyk, P., & Luce, P. (2002). Lexical neighbourhood effects in 17-month-old word learning. *Proceedings of the 26th Annual Boston University Conference on Language Development* (pp. 314–23). Boston, MA: Cascadilla Press.

Kuhl, P. K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show facilitation for native language phonetic perception between 6 and 12 months. *Developmental Science*, *9*, 13–21.

Lacerda, F. (1993). Sonority contrasts dominate young infants' vowel perception. *Journal of the Acoustical Society of America*, *93*, 2372.

Lacerda, F. (1994). The asymmetric structure of the infant's perceptual vowel space. *Journal of the Acoustical Society of America*, *95*, 3016.

MacKenzie, H., Curtin, S., & Graham, S. A. (2012). 12-month-olds' phonotactic knowledge guides their word–object mapping. *Child Development*, *83*, 1129–36.

MacKenzie, H., Graham, S. A., & Curtin, S. (2011). 12-month-olds privilege words over other linguistic sounds in an associative learning task. *Developmental Science*, *14*, 249–55.

MacKenzie, H., Graham, S. A., Curtin, S., & Archer, S. L. (2014). The flexibility of 12-month-olds' preferences for phonologically appropriate object labels. *Developmental Psychology*, *50*(2), 422–30.

May, L., & Werker, J. F. (2014). Can a click be a word? Infants' learning of non-native words. *Infancy*, *19*(3), 281–300.

Narayan, C. R., Werker, J. F., & Beddor, P. S. (2010). The interaction between acoustic salience and language experience in developmental speech perception: evidence from nasal place discrimination. *Developmental Science*, *13*(3), 407–20.

Pater, J., Stager, C., & Werker, J. F. (2004). The perceptual acquisition of phonological contrasts. *Language*, *80*, 384–402.

Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(2), 421–35.

Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word learning tasks. *Nature*, *388*, 381–2.

Storkel, H., & Rogers, M. (2000). The effect of probabilistic phonotactics on lexical acquisition. *Clinical Linguistics and Phonetics*, *14*, 407–25.

Thiessen, E. D. (2007). The effect of distributional information on children's use of phonemic contrasts. *Journal of Memory and Language*, *56*, 16–34.

Thiessen, E. D., & Yee, M. N. (2010). Dogs, bogs, labs, and lads: what phonemic generalizations indicate about the nature of children's early word-form representations. *Child Development*, *81*(4), 1287–303.

Wang, Y., & Seidl, A. (2015). The learnability of phonotactic patterns in onset and coda positions. *Language Learning and Development*, *11*(1), 1–17.

Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word–object associations by 14-month old infants. *Developmental Psychology*, *34*, 1289–309.

Werker, J. F. & Curtin, S. (2005). PRIMIR: a developmental framework of infant speech processing. *Language Learning and Development*, *1*, 197–234.

Werker, J. F., Fennell, C. T., Corcoran, K. M., & Stager, C. (2002). Infants' ability to learn phonetically similar words: effects of age and vocabulary size. *Infancy*, *3*, 1–30.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *7*, 49–63.

Yoshida, K. A., Fennell, C. T., & Swingley, D. (2009). Fourteen-month-old infants learn similar-sounding words. *Developmental Science*, *12*, 412–18.

Zamuner, T. S. (2006). Sensitivity to word-final phonotactics in 9- to 16-month-old infants. *Infancy*, *10*, 77–95.

Appendix A

Detailed report of participant data not included in the analysis.

	Experiment 1	Experiment 2	Experiment 3
Parental distraction	3		4
Parent stops experiment	2	2	1
Fussy/cried/distracted[a]	6	6	
Technical/experimenter errors	1		2
Did not habituate	1	1	1
Eyes obscured[b]	2	1	
2.5 Standard deviations from mean		1	
Looking <1 second		1	
Total	15	12	8

Notes. [a] = Eye gaze unreadable during offline coding, crying loudly drowning out audio, and/or consistently distracted by other objects than those on screen; [b] = Infant's eyes outside of camera frame.

Appendix B

Interaction results of 2 × 2 ANOVAs between Experiments 1-3 and 2-3

Exp. 1 – Exp. 3		
2 × 2 ANOVA	*Interaction (n.s.)*	$F(1,30) = .511, p = .480$
Exp. 2 – Exp. 3		
2 × 2 ANOVA	*Interaction (n.s.)*	$F(1,30) = 5.532, p = .313$

Means and Standard Deviations for test trials by Experiment

Experiment	Test Trial	Mean (*M*), Standard Deviation (*SD*)	Pairwise t-tests
Exp. 1	same	$M = 6.83, SD = 3.95$	$t(15) = −2.253, p = .040, d = .594$
	switch	$M = 8.83, SD = 4.67$	
Exp. 2	same	$M = 6.25, SD = 3.16$	$t(15) = −3.146, p = .007, d = .787$
	switch	$M = 8.63, SD = 4.56$	
Exp. 3	same	$M = 8.00, SD = 5.22$	$t(15) = −.785, p = 444$
	switch	$M = 8.93, SD = 4.70$	