Developmental Science

Developmental Science DOI: 10.1111/j.1467-7687.2009.00814.x

PAPER

Weighting of vowel cues explains patterns of word–object associative learning

Suzanne Curtin,[1] Christopher Fennell[2] and Paola Escudero[3]

1. Departments of Linguistics and Psychology, University of Calgary, Canada
2. School of Psychology and the Department of Linguistics, University of Ottawa, Canada
3. Institute of Phonetic Sciences, University of Amsterdam, The Netherlands

Abstract

Previous research has demonstrated that infants under 17 months have difficulty learning novel words in the laboratory when the words differ by only one consonant sound, irrespective of the magnitude of that difference. The current study explored whether 15-month-old infants can learn novel words that differ in only one vowel sound. The rich acoustic/phonetic properties of vowels allow for a detailed analysis of the contribution of acoustic/phonetic cues to infants' performance with similar-sounding words. Infants succeeded with the vowel pair /i/–/I/, but failed with vowel pairs /i/–/u/ and /I/–/u/. These results suggest that infants initially use the most salient acoustic cues for vowels and that this staged use of acoustic cues both predicts and explains why infants can learn some words that differ in only a single vowel.

Introduction

Around their first birthday, infants detect subtle mispronunciations of known object labels (Bailey & Plunkett, 2002; Fennell & Werker 2003; Swingley, 2005; Swingley & Aslin, 2002), suggesting that their lexical representations encode a great deal of phonetic information. However, studies using novel minimally different words yield different results. When learning two novel word–object pairings where the words differ by only one consonant, infants under 17 months generally have difficulty detecting a mismatch in the pairing, demonstrating that they are not accessing all the available information (Pater, Stager & Werker, 2004; Stager & Werker, 1997; Werker, Fennell, Corcoran & Stager, 2002; but see Ballem & Plunkett, 2005, and Fennell, 2006, for success under enriched learning conditions). Importantly, this is not a general problem with word learning, as same-age infants can learn novel pairings when the words are maximally different, like 'lif' and 'neem' (Stager & Werker, 1997).

To date, work with similar-sounding words has predominantly focused on consonant differences. Infants' use of vowel detail may resemble that of consonants, with less robust use at younger ages, or it may emerge in a different manner, either appearing developmentally earlier or later than consonants. One argument for an earlier emergence of vowel use in word learning relates to the fact that infants narrow their perceptual focus to native vowels by 6 months of age (Kuhl, Williams, Lacerda, Stevens & Lindblom, 1992; Polka & Werker, 1994), several months before narrowing their perceptual focus to native consonants (Werker & Tees, 1984). In contrast, Nespor, Peña and Mehler (2003) argue that infants should have greater focus on consonants than on vowels in early word learning since vowels are more susceptible to changes in pronunciation and are perceived less categorically (Pisoni, 1973).

Very few studies have specifically compared infants' use of vowel versus consonant information in early words. Mani and Plunkett (2007) found that 18-month-olds recognize mispronunciations of word-medial vowels in familiar words, but this finding is less robust at 15 months, whereas infants of both ages notice consonant mispronunciations. This pattern of results may indicate that vowels are not used as efficiently as consonants in early word learning. Similarly, results from interactive object categorization tasks with novel words demonstrated that 20-month-olds confuse labels that minimally differ in their vowel, but succeed in distinguishing labels that minimally differ in consonants (Nazzi, 2005; Nazzi & New, 2007). Based on these studies, it appears that infants perform worse on vowels than consonants, as predicted by Nespor *et al.* (2003). However, 18-month-olds appear to be sensitive to small changes to vowels in familiar (Mani, Coleman & Plunkett, 2008) and novel words (Dietrich, Swingley & Werker, 2007), and 14-month-olds can notice broad vowel mispronunciations (i.e. three feature changes) in recently learned novel words (Mani & Plunkett, 2008). Thus, the picture relating to novice

Address for correspondence: Suzanne Curtin, Department of Psychology, 2500 University Drive NW, University of Calgary, Calgary, AB Canada T2N 1N4; e-mail: scurtin@ucalgary.ca

word learners' use of vowel information in words is not as clear as with consonants, perhaps due to the use of different tasks, vowel contrasts (including multiple-feature contrasts), and different word types (familiar versus novel).

To obtain a clearer picture of vowel use in early word learning, we tested infants' ability to detect close vowel pairs (/i/–/I/, /i/–/u/, and /I/–/u/) in a novel word–object pairing task used in previous research with consonants: the Switch procedure (Werker *et al.*, 2002). The vowel /i/ is phonologically characterized as a high, front, tense vowel. The vowel /I/ is a mid-high, front, lax vowel, and /u/ is a high, back, round, tense vowel. The tense/lax distinction will not be further considered in this paper, as researchers have argued that 'lax' vowels are centralized, and the degree of frontness, which acoustically maps onto vowel formants, is a better characterization of these vowels (Ladefoged, 2007). Vowel formants reflect the resonant frequencies of the vocal tract and provide information about vowel identity. The first formant (F1) corresponds to height, the second formant (F2) to backness, and the third formant (F3) to lip rounding. Front 'lax' vowels (e.g. [I]) are lower (higher F1) than front 'tense' vowels, (e.g. [i]) and more central (lower F2) (Ladefoged & Maddieson, 1996). In sum, /i/ and /I/ differ in height and backness (F1 and F2), /i/ and /u/ differ in backness and rounding (F2 and F3), and /I/ and /u/ differ in height, backness and rounding (F1, F2, and F3).

A number of predictions surface depending on what information infants store when learning new words and whether or not they treat vowels as less important for lexical items than consonants. If infants store all information about word–object pairings, they should succeed on all contrasts; a possibility supported by infants' earlier perceptual refinement of vowels. However, considering their difficulties with consonantal information at this age (e.g. Stager & Werker, 1997) and with vowel information at later ages (e.g. Nazzi, 2005), this possibility may be unlikely. Thus, the question of interest becomes: will infants succeed on all, some, or none of the vowel contrasts? Three theoretical perspectives have differing predictions on this matter.

According to the developmental framework PRIMIR (Processing Rich Information from Multidimensional Interactive Representations), phonetic category formation takes place concurrently with the formation of the lexicon (Werker & Curtin, 2005). Access to information stored in representations depends on the task an infant is confronted with at a particular age, the salience and frequency of the information, and the current state of organization of the representations. This dynamic system encodes all types of linguistic information, but access depends on the current state of the system. If infants do not yet have phoneme representations, which PRIMIR argues is the case for younger word learners such as those in the current study, they must rely on other relevant sources of information, such as acoustic/phonetic cues. Since the system is dynamic and differential weighting of cues occurs as words are stored and phonetic categories are emerging, infants may rely on some acoustic features more so than others. However, it is not clear which acoustic features (i.e. formants) might be more heavily weighted.[1] The current results may shed light on these weightings if infants perform differently on the contrasts. Thus, PRIMIR predicts that infants should succeed on some or all of the vowel contrasts, considering that vowels have more salient acoustic information (longer and louder) associated with them than the consonants used in past research with infants this age.

Alternatively, some researchers argue that stored *phonological* representations in the lexicon are underspecified (Fikkert, in press; Pater *et al.*, 2004). If so, then infants will not have the necessary information to distinguish between certain contrasts. In particular, Fikkert argues that coronals are lexically underspecified (see also Lahiri & Reetz, 2002). When the child encounters a word that contains a coronal, the representation itself does not contain any featural information. Support for this prediction has been found with consonants at 14 months of age (Fikkert, in press).[2] In our study, all the consonants are coronal, as are our front vowels (/i/ and /I/), whereas the place of articulation for the back vowel /u/ is dorsal (and possibly labial if rounding is contributing). Thus, aside from the specified place information in /u/, the only information available in the vowels is height. The vowels /i/ and /u/ are high and /I/ is mid-high. If coronal is underspecified, infants should succeed on /dit/–/dIt/ and /dIt/–/dut/ if they are relying on just the feature height. The predictions are unclear for /dit/–/dut/. While /u/ is dorsal, /i/ has no features to contrast with /u/ and both vowels are high, so infants may not have the necessary contrastive information to distinguish between the words (see also Escudero & Benders, in press).

Finally, if young word learners weigh vowel information less than consonant information in word recognition, as per Nespor *et al.*'s (2003) model, they should not succeed on any of these contrasts as all novel words in this study have the same consonants. The results of this study will not be able to directly address the question of whether the role of vowels versus consonants in lexical access changes over the course of development as we are not directly comparing the two in the current experiment. However, past research has clearly demonstrated that infants this age fail with consonants that differ in up to two features (Pater *et al.*, 2004). Therefore, if infants succeed on any of the vowel contrasts, especially those differing in two

---

[1] The formal linguistic models of Boersma, Escudero and Hayes (2003) and Escudero (2005) argue that perceptual representations are copied to the lexicon with their specific level of abstraction at a specific time in development. It is also overtly proposed that young infants have one-dimensional, abstract perceptual categories (i.e. height or backness) and that the weighting of perceptual cues depends on their reliability in the input.

[2] Fikkert (in press) habituated infants to only a single object. As a result her predictions result in asymmetries. We, however, taught infants two novel word–object pairings, which complicate her model's predictions.

Table 1 *Acoustic measurements: averages across the 10 tokens per vowel (standard deviations are in parentheses)*

Vowel	F1 Hz	F2 Hz	F3 Hz	Duration ms
/i/	389.1 (44.8)	2622.2 (121.5)	3025.5 (182.5)	302.6 (42.1)
/I/	620.2 (73.4)	2276.8 (111.1)	2937.8 (158.7)	245.7 (28.7)
/u/	451.4 (42.2)	1496.2 (114.7)	2471.8 (199.6)	300.8 (38.5)

features, it would indicate that vowels can be used in word learning at a time when consonants cannot be used efficiently, thus posing a potential issue for this model that would require further investigation.

Experiment

We tested whether 15-month-olds could detect a mismatch in the word–object pairing when the words nominally differ in their vowels. Three pairings were tested: ‘deet’–‘dit’, ‘deet’–‘doot’, and ‘dit’–‘doot’.

Method

Participants

Fifty-four 15-month-olds participated in this study. Sixteen infants (8 males) were presented novel words that differed in the vowels /i/ and /I/ (i.e. ‘deet’ and ‘dit’). Eighteen infants (10 males) heard words differing in the vowels /i/ and /u/ (i.e. ‘deet’ and ‘doot’). Twenty infants (10 males) heard words differing in the vowels /I/ and /u/ (i.e. ‘dit’ and ‘doot’). The mean age of the participants was 15 months, 15 days (14 m, 21 d–16 m, 5 d). Twenty-nine additional infants were tested but were excluded from the analysis because they were fussy (11), there was interference from the parent (12), they did not habituate (4), or they were too restless to code properly (2).

Stimuli

A female native speaker of Canadian English produced the three CVC non-words. Stimuli were recorded at a 44 kHz sample rate directly onto a Macintosh G3 computer using Sound Edit 16 (Version 2-99) software. The speaker produced each CVC item 10 times using a range of infant-directed contours. Vowel duration and formants were measured using Praat 4.2 (Boersma & Weenink, 2004). All measurements were taken from the mid-point of the vowel (50% of total vowel length).

Visual stimuli used in the habituation and test phases consisted of two attractive novel objects (see Figure 1a and b). A toy waterwheel was used for both the pre- and post-tests (see Figure 1c). The habituation and test objects were presented moving back and forth across the screen at a slow and constant speed. The waterwheel was filmed stationary with its arms moving in a rotating motion.

Apparatus

Participants were tested in a dimly lit, sound attenuated, small room. Infants sat on their parents’ laps facing a display (a 27-inch Mitsubishi CS 27205C video monitor in the /i/–/I/ condition and a 122-cm high by 91.5-cm wide video monitor in the other conditions). Audio stimuli were delivered at 65 dB, ±5 dB. Infants were recorded with a digital video camera for subsequent frame-by-frame coding. The Habit 2000 program (Cohen, Atkinson & Chaput, 2000) was used to order stimuli presentation and collect looking time data for the /I/–/i/ condition. Habit X 1.0 (Cohen, Atkinson & Chaput, 2004) was used for the other conditions.

Procedure

We used a Switch design that matched Werker *et al.* (2002). During habituation, infants were presented with two word–object pairings. One pairing consisted of the crown object paired with the label ‘dit’ in the first and third conditions and ‘doot’ in the second (Pair A). The other pairing was the molecule object paired with the label ‘deet’ in the first two conditions and ‘doot’ in the third (Pair B). In each trial, infants heard all ten tokens of the relevant word. Each randomized block included two trials of each pairing type, resulting in six possible pairing orders within a block (e.g. ABAB, BAAB, etc.). Once the habituation criterion was met (65% of looking time from the first block), two test trials were presented. In the ‘Same’ trial, a word and an object were presented in the appropriate combination (e.g. Object A–Word A). In the ‘Switch’ trial, the pairing was violated (e.g. Object A–Word B). If infants learned about the words and the objects, but did not learn the crucial associative link, the ‘Same’ and ‘Switch’ trials would be equally familiar, resulting in equal looking times for both trials. If infants learned the link between the specific words and objects, they should look longer to the ‘Switch’ than to the ‘Same’ trial.

The waterwheel–‘pok’ combination, a completely novel word–object pairing, was presented after the test trials to

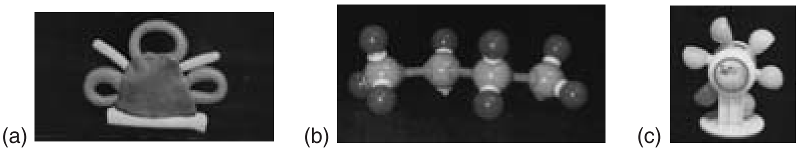

Figure 1 *Visual stimuli used in the experiment.*

ensure that the infants recovered to a large change, and were not fatigued or generally disinterested in the task. This combination also preceded the habituation phase as a pre-test.

Results

A 2 (test trials: Same vs. Switch) × 3 (condition: /i/–/I/ vs. /i/–/u/ vs. /I/–/u/) mixed ANOVA revealed no significant main effects for the within-subjects factor of test trial [$F(1, 51) = 1.93, p = .17, \eta_p^2 = .04$], or for the between-subjects factor of condition [$F(2, 51) = .74, p = .48, \eta_p^2 = .03$]. However, the interaction was significant and accounted for the largest amount of variance, as indicated by the effect size [$F(2, 51) = 3.2, p = .049, \eta_p^2 = .11$]. This led to a closer analysis of the test trial performance across conditions. All looking times are presented in seconds. Infants in the /i/–/I/ condition looked significantly longer to Switch ($M = 10.85$; $SD = 5.46$) than to Same ($M = 7.25$; $SD = 4.39$) trials [$t(15) = -2.23, p = .04, d = -.73$]. There was no significant difference between Switch ($M = 7.69$; $SD = 5.47$) and Same ($M = 8.33$; $SD = 4.44$) in the /i/–/u/ condition [$t(17) = .55, p = .59, d = .13$] or between Switch ($M = 7.54$; $SD = 3.82$) and Same ($M = 7.53$; $SD = 3.81$) in the /I/–/u/ condition [$t(19) = -.01, p = .99, d = -.003$] (see Figure 2). These results were further confirmed by chi-squared analyses investigating the number of infants who looked longer to the Switch over the Same trial in each condition. In the /i/–/I/ condition, 12 of the 16 infants looked longer to Switch than to Same, exceeding chance [$\chi^2(1) = 4.0, p = .046$]. The number of infants looking longer to the Switch trial did not exceed chance in the two other conditions: 8 of 18 infants in the /i/–/u/ condition [$\chi^2(1) = .22, p = .64$] and 11 of 20 infants in the /I/–/u/ condition [$\chi^2(1) = .2, p = .65$].

To ensure that infants were not generally fatigued or uninvolved in the task, we compared the post-test trial to the average looking time in the last block of habituation stimuli. A mixed 2 (trials: post-test versus last habituation block) by 3 (vowel: /i/–/I/, /i/–/u/, /u/–/I/) ANOVA revealed a significant effect of trials [$F(1, 51) = 440.87, p < .001; \eta_p^2 = .90$], with the infants looking longer to the post-test ($M = 17.44$, $SD = 3.36$) than to the last block of habituation trials ($M = 7.81$, $SD = 3.09$). There was no interaction between vowels and trials [$F(2, 51) = .92, p = .41; \eta_p^2 = .04$]. These findings demonstrate that infants were still engaged in the task across all vowel conditions. There was a main effect for vowel condition trials [$F(12, 51) = 4.26, p = .02; \eta_p^2 = .04$]. Post-hoc *t*-tests revealed that infants in the /u/–/I/ condition had significantly lower looking times across the two trial types ($M = 11.34$; $SD = 3.00$) than infants in the /i/–/I/ condition ($M = 13.80$; $SD = 1.51$), whereas the infants' mean looking time across the trials in the /i/–/u/ condition ($M = 13.00$; $SD = 2.87$) did not differ from either of the other conditions. This may indicate that infants in the /u/–/I/ condition were overall less interested overall in the task than the infants in the /i/–/I/ condition, but their significant recovery to the post-test allows for a valid interpretation of their data.

Finally, we examined whether the number of habituation trials influenced infant success. Infants experienced between 8 and 24 habituation trials, but this number varied across infants based on when the criterion was met. Using a one-way ANOVA, we found that the number of habituation trials significantly differed across conditions [$F(2, 51) = 5.23, p = .009$]. Post-hoc analyses revealed that the only significant difference is between Deet–Doot and Doot–Dit (Deet–Dit: $M = 13.2$, $SD = 3.79$; Deet–Doot: $M = 16.67$, $SD = 5.18$; Doot–Dit: $M = 11.80$, $SD = 4.94$). Therefore, the condition with the highest number of habituation trials and that with the lowest were both unsuccessful conditions. While this is interesting, it does not suggest any relationship between success and habituation trials.

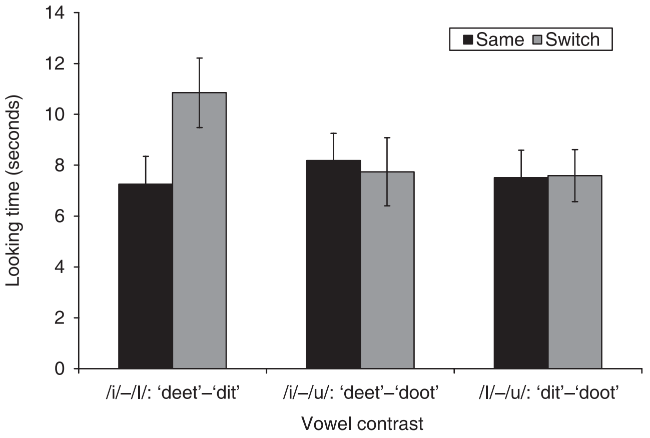

Figure 2 *Results for all vowel contrasts.*

Discriminant analysis

The 15-month-old infants' longer looking time to Switch in the case of the /i/–/I/ contrast demonstrates that novice word learners can use at least some detailed vowel differences to direct word learning. However, their failure with the other two contrasts suggests that not all vowels are treated similarly by infants. We ran a discriminant function analysis on the 30 tokens with which infants were presented (10 tokens per vowel). Discriminant function analysis (DA) is a statistical procedure that classifies items into categories based on a set of independent variables. The DA was not run on each contrast separately because we assume that infants are classifying more than one vowel at a time. From this model, contrast-based inferences can be made with the assumption that all three vowels could be confused with one another.

The DA was conducted to see whether four acoustic predictors – F1 (height), F2 (backness), F3 (rounding), and duration[3] – could identify the three vowel categories.

---

[3] The vowel /I/ is acoustically shorter than both /i/ and /u/. This difference is not considered to be part of the phonological feature specifications for English vowels since this is not a contrastive feature for the language. However, it is possible that infants are sensitive to these differences.

Table 2 *Confusion matrix for discriminant analysis model*

Tokens (N = 10)	Vowels' F1 or F2 values only					
	/i/ (F1)	/I/ (F1)	/u/ (F1)	/i/ (F2)	/I/ (F2)	/u/ (F2)
/i/ (production)	10	0	0	8	**2**	0
/I/ (production)	0	10	0	0	10	0
/u/ (production)	**2**	**2**	6	0	0	10

Notes: Bold = Misclassifications; F2 dimension alone: 93.33% accuracy; F1 dimension alone: 86.66% accuracy.

Tokens (N = 10)	Vowels' F3 or duration values only					
	/i/ (F3)	/I/ (F3)	/u/ (F3)	/i/ (dur.)	/I/ (dur.)	/u/ (dur.)
/i/ (production)	6	**4**	0	4	**2**	**4**
/I/ (production)	**3**	7	0	**1**	8	**1**
/u/ (production)	0	0	10	**6**	**3**	1

Bold = Misclassifications; F3 dimension alone: 76.66% accuracy; Duration dimension alone: 43.33% accuracy.

The overall Wilks' lambda was significant (Λ = .01, χ^2(4, N = 24) = 100.75, p < .01), indicating that all predictors combined differentiated the three vowels. Thus, when all cues are present, the three vowels are reliably classified.

To assess the classification power of each individual acoustic dimension, we used a classification model that assigns 30 tokens of the three vowels to categories purely on the basis of either their F1, F2, F3, or duration values. Each model was trained with eight tokens per vowel and was then required to classify the 24 tokens with which the model was trained for cross-validation. As a final step, the model classified six tokens (two tokens per vowel) that were not included in the training.

According to the results, duration was the least reliable predictor for any of the vowel contrasts. F3 led to good classification for /i/–/u/. However, infants had difficulty with this contrast, suggesting that F3 is not being used efficiently at this age. In the case of F2, there was confusion for /i/–/I/ but good classification for /i/–/u/ and /u/–/I/. Again, our results suggest that infants do not exploit F2 differences because they do not succeed with the /i/–/u/ contrast. There was also confusion for /i/–/u/ and /I/–/u/ but good classification for /i/–/I/ along the F1 dimension. This suggests that infants exploited F1 differences because they succeeded only with the /i/–/I/ contrast, the only contrast which exhibited no overlap along the F1 dimension. These findings suggest that infants rely on the relative importance of the height dimension.

General discussion

Infants can detect a switch in a word–object pairing that differs in only a vowel, but only for one of three vowel contrasts tested. If we rely purely on a level of description which weights any change in vowel equally (i.e. an adult-like phonological feature system), then infants should have been successful on all tested contrasts. Similarly, if word learners did not consider vowel information to be as important in lexical items as consonants, as postulated by Nespor *et al.* (2003), they should have failed in all three vowel conditions. However, 15-month-olds were successful on the /i/–/I/ contrast. This does not follow Fikkert's underpsecification model either, as it predicts that infants would succeed on the /I/–/u/ contrast as well. Thus, it appears that infants are weighting certain cues more than others, a possibility allowed for in PRIMIR.

Formal phonological modeling predicts that infants are able to use cues separately according to the reliability in the input (Boersma, Escudero & Hayes, 2003; Escudero, 2005). When we examine the acoustic properties of all the vowels used in this study, we see that categorization accuracy based on F1 (height) is 100% for /i/–/I/. This was not the case with the /i/–/u/ contrast or the /I/–/u/ contrast – both of which differed most along the F2 (backness) dimension, suggesting that early in word learning infants rely more on F1 to distinguish between vowel contrasts. This opens up the possibility that infants will succeed with other contrasts that differ only in F1. Mani *et al.* (2008) found 18-month-olds were more sensitive to changes in vowel height and backness and less so for rounding. Perhaps the weighting of F2 increases as infants' repertoire of meaningful sound categories grows along with the development of their lexicon.

There are several possible explanations for these results based on infants' linguistic environments. F1 is more salient than F2 because there is more acoustic energy in F1 (Lacerda, 1993, 1994). Infants' reliance on F1 in this task may also be explained by the greater number of height over backness category distinctions in English vowels. Additionally, there is strong evidence that the back vowel /u/ is quite fronted in North American English (Thomas, 2001). This general instability of /u/ might decrease F2's, and even F3's, reliability as a cue. Further studies examining F1 and F2 relationships between different vowels, dialects, and the role of salience will help to distinguish these possible explanations.

Early on, infants appear to use reliable auditory differences along individual continua, like F1. However, once infants have a phonological system in place, phonemic differences rather than acoustic salience should guide word learning. The emergence of phonemes is supported by a cross-linguistic study examining the use of vowel duration by 18-month-old infants (Dietrich *et al.*, 2007). Dutch-learning infants, for whom vowel duration is phonological, successfully distinguished novel words that differed in the duration of their medial vowel, whereas English-learning infants, for whom vowel duration serves no phonological purpose, failed. This suggests, as predicted by PRIMIR, that phonemes emerge over time and become weighted more than individual acoustic cues.

In conflict with our results and those of Dietrich *et al.* (2007) is the failure of 20-month-old infants to use vowel detail in Nazzi (2005). Perhaps Nazzi's object categorization task placed different demands on infants than the object–label associative task used in both our study and Dietrich *et al.*'s work. Infants may relax their tolerance for vowel differences when dealing with a real speaker and multiple labels in the interactive categorization task. Infants in our task hear only one speaker producing the words and the words are heard repeatedly during habituation. Vowel differences, once developmentally available, may be highlighted in attention due to our design. Moreover, it may be that as the lexicon expands, the roles and weightings of consonants and vowels as cues to lexical distinctiveness change over time (Nespor *et al.*, 2003). Indeed, our results suggest that early on salience of acoustic information informs lexical access, a finer level than the broad vowel versus consonant distinction. Further research using both tasks and comparing different vowel contrasts is needed to determine the developmental trajectory of vowel use over infancy.

To date, our study is one of the first to demonstrate that infants under 17 months of age detect vowel contrasts in *two newly learned* word–object pairings (see also Mani & Plunkett, 2008) and the first to systemically test multiple vowel contrasts at this age. Importantly, our findings demonstrate that this ability does not generalize to all vowels. Fifteen-month-olds did not look longer at a mismatch in the pairing for all vowel contrasts. Indeed, they do so only for the contrast that differs along the F1 dimension (vowel height). These findings demonstrate the role of salience and the weighting of acoustic properties in the recognition of newly learned word–object pairings, and further characterizes the developmental trajectory of early word learning with respect to consonant and vowel contrasts.

Acknowledgements

We thank Danielle Droucker and Natasha Nickel for their assistance with this research. We especially thank Janet Werker for her extremely helpful discussions and comments on this research, and for her generous funding support (NSERC Discovery Grant A1103). We would also like to thank Denis Mareschal and two anonymous reviewers for their constructive and insightful comments. We also thank the families who participated in this research. This research was also supported by an SSHRC post-doctoral grant (No. 756-2001-43) and a URGC grant through the University of Calgary awarded to Suzanne Curtin. Paola Escudero's work was supported by a Netherlands Organization for Scientific Research (NWO) grant (No. 016.024.018) awarded to Paul Boersma.

References

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

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

Boersma, P., Escudero, P., & Hayes, R. (2003). Learning abstract phonological from auditory phonetic categories: an integrated model for the acquisition of language-specific sound categories. In M.J. Sole, D. Recansens, & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences*, Barcelona.

Boersma, P., & Weenink, D. (2005). *Praat: doing phonetics by computer* (Version 4.5.16) [Computer program]. Downloaded from http://www.praat.org/.

Cohen, L.B., Atkinson, D.J., & Chaput, H.H. (2000). *Habit 2000: A new program for testing infant perception and cognition* (Version 2.2.5c) [Computer software]. Austin, TX: University of Texas.

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, TX: University of Texas.

Dietrich, C., Swingley, D., & Werker, J.F. (2007). One-year-olds' language-specific phonological categorization in word learning: a cross-linguistic study. *Proceedings of the National Academy of Sciences of the USA*, **104** (41), 16027–16031.

Escudero, P. (2005). Linguistic perception and second language acquisition: explaining the attainment of optimal phonological categorization. Doctoral dissertation, Utrecht University. LOT Dissertation Series 113.

Escudero, P., & Benders, T. (in press). Phonetic and phonological approaches to early word recognition: empirical findings, methodological issues, and theoretical implications. In T. Lentz, H. De Mulder, Ø. Nilsen, & A. Zondervan (Eds.), *Psychological reality and theoretical validity*. Amsterdam: John Benjamins.

Feldman, H.M., Dollaghan, C.A., Campbell, T.F., Kurs-Lasky, M., Janosky, J.E., & Paradise, J.L. (2000). Measurement properties of the MacArthur Communicative Development Inventories at ages one and two years. *Child Development*, **71** (2), 310–322.

Fennell, C.T. (2006). Infants of 14 months use phonetic detail in novel words embedded in naming phrases. *Proceedings of the 30th Annual Boston University Conference on Language Development*, 178–189.

Fennell, C.T., & Werker, J.F. (2003). Early word learners' ability to access phonetic detail in well-known words. *Language and Speech*, **46** (2), 245–264.

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

Jusczyk, P.W. (1997). *The discovery of spoken language*. Cambridge, MA: MIT Press.

Kuhl, P.K. (1987). The special-mechanisms debate in speech research: categorization tests on animals and infants. In S. Harnad (Ed.), *Categorical perception: The groundwork of cognition* (pp. 355–386). New York: Cambridge University Press.

Kuhl, P.K., Williams, K.A., Lacerda, F., Stevens, K.N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, **255**, 606–608.

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.

Ladefoged, P. (2007). Articulatory features for describing lexical distinctions. *Language*, **83** (1), 161–180.

Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Oxford: Blackwell.

Lahiri, A., & Reetz, H. (2002). Underspecified recognition. In C. Gussenhoven, Natasha Werner, & Toni Rietveld (Eds.), *Labphon 7* (pp. 637–676). Berlin: Mouton.

Mani, N., Coleman, J., & Plunkett, K. (2008). Phonological specificity of vocal features at 18-months. *Language and Speech*, **51**, 3–21.

Mani, N., & Plunkett, K. (2007). Phonological specificity of vowels and consonants in early lexical representations. *Journal of Memory and Language*, **57**, 252–272.

Mani, N., & Plunkett, K. (2008). Fourteen-month-olds pay attention to vowels in novel words. *Developmental Science*, **11** (1), 53–59.

Nazzi, T. (2005) Use of phonetic specificity during the acquisition of new words: differences between consonants and vowels. *Cognition*, **98** (1), 13–30.

Nazzi, T., & New, B. (2007). Beyond stop consonants: consonantal specificity in early lexical acquisition *Cognitive Development*, **22** (2), 271–279.

Nespor, M., Peña, M., & Mehler, J. (2003). On the different roles of vowels and consonants in speech processing and language acquisition. *Lingue e Linguaggio*, **ii**, 221–247.

Pater, J., Stager, C.L., & Werker, J.F. (2004). The lexical acquisition of phonological contrasts. *Language*, **80** (3), 361–379.

Pisoni, D.B. (1973). Auditory and phonetic memory codes in the discrimination of consonants and vowels. *Perception and Psychophysics*, **13**, 253–260.

Polka, L., & Werker, J.F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, **20** (2), 421–435.

Stager, C.L., & Werker, J.F. (1997). Infants listen for more phonetic detail in speech perception than in word learning tasks. *Nature*, **388**, 381–382.

Swingley, D. (2005). Eleven-month-olds' knowledge of how familiar words sound. *Developmental Science*, **8**, 432–443.

Swingley, D., & Aslin, R.N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, **13**, 480–484.

Thomas, E.R. (2001). *An acoustic analysis of vowel variation in New World English*. Durham, NC: Duke University Press.

Tincoff, R., & Jusczyk, P.W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science*, **10** (2), 172–175.

Werker, J.F., Cohen, L.B., Lloyd, V., Stager, C.L., & Cassasola, M. (1998). Acquisition of word–object associations by 14-month-old infants. *Developmental Psychology*, **34** (6), 1289–1309.

Werker, J.F., & Curtin, S. (2005). PRIMIR: a developmental framework of infant speech processing. *Language Learning and Development*, **1** (2), 197–234.

Werker, J.F., Fennell, C.T., Corcoran, K., & Stager, C.L. (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.

Received: 12 October 2007
Accepted: 11 June 2008