Developmental Science

Developmental Science 12:3 (2009), pp 412–418 DOI: 10.1111/j.1467-7687.2008.00789.x

PAPER

Fourteen-month-old infants learn similar-sounding words

Katherine A. Yoshida,[1] Christopher T. Fennell,[2] Daniel Swingley[3] and Janet F. Werker[1]

1. Department of Psychology, University of British Columbia, Canada
2. School of Psychology, University of Ottawa, Canada
3. Department of Psychology, University of Pennsylvania, USA

Abstract

Can infants, in the very first stages of word learning, use their perceptual sensitivity to the phonetics of speech while learning words? Research to date suggests that infants of 14 months cannot learn two similar-sounding words unless there is substantial contextual support. The current experiment advances our understanding of this failure by testing whether the source of infants' difficulty lies in the learning or testing phase. Infants were taught to associate two similar-sounding words with two different objects, and tested using a visual choice method rather than the standard Switch task. The results reveal that 14-month-olds are capable of learning and mapping two similar-sounding labels; they can apply phonetic detail in new words. The findings are discussed in relation to infants' concurrent failure, and the developmental transition to success, in the Switch task.

Introduction

Over the second half of the first year of life, infants begin to acquire words, first as sequences of sounds, and eventually as meaningful words in the developing vocabulary (Jusczyk & Hohne, 1997; Swingley, 2007). In the same developmental period speech perception is refined, with speech sound discrimination declining for non-native language sound contrasts (Werker & Tees, 1999), improving for native contrasts (Kuhl, Stevens, Hayashi, Deguchi, Kiritani & Iverson, 2006; Narayan, 2006), and realigning to native phonetic boundaries (Burns, Yoshida, Hill & Werker, 2007). The phonetic and lexical developments that take place in the first year indicate a surprising early capacity for learning from the speech signal. Given the cognitive skills required for discovering a language's system of phonetic categories, it would seem that the application of these talents to the problem of word learning in the second year would be a simple matter, and certainly not a limiting factor in building the lexicon. Yet the literature on phonological encoding in 1-year-olds' word learning does not paint a simple picture. Indeed, although infants can distinguish known words from close approximations (Swingley & Aslin, 2002; see also Mani & Plunkett, 2007; Swingley, 2005; White & Morgan, 2008), mapping two similar-sounding words to referents in word learning tasks appears more difficult (Pater, Stager & Werker, 2004; Stager & Werker, 1997). The discrepancy in findings may be due in part to the range of methodologies used to test the question (Werker & Curtin, 2005).

One of these methods is the audiovisual habituation Switch task, in which infants are presented with two word–object pairings (object-a is presented with word label-a and object-b is presented with label-b) until looking time drops below a preset criterion. Two test trials assess whether infants have associated the words with their appropriate objects. The control 'same' trial presents a now-familiar combination (object-a with label-a) and the 'Switch' trial presents a familiar word and object in a new combination (object-a with label-b). If the infants have not successfully formed the associations, the sounds and objects in the same and Switch trials will be equally familiar and the trials should attract comparable looking times. However, if infants have tied the labels to the associated objects, the Switch trial should surprisingly violate the linkage, resulting in greater looking (Werker, Cohen, Lloyd, Casasola & Stager, 1998). To date, the results from this procedure suggest that although infants of 14 months succeed in using discriminable phonetic detail (i.e. increased looking to the Switch trial) with familiar words (*ball* and *doll*; Fennell & Werker, 2003), they fail with new words (e.g. *bih* and *dih*; Stager & Werker, 1997), only succeeding by 17 months of age (Werker, Fennell, Corcoran & Stager, 2002).

Variants on the Switch task that augment the nature of the learning experience improve 14-month-olds' performance with new words. For example, infants are more successful in noticing changes in the pronunciations of new words if they are prefamiliarized to the objects (Fennell, 2004), if the words are embedded in simple naming phrases (Fennell, 2006), if a training phase is

Address for correspondence: Janet Werker, 2136 West Mall, Vancouver, BC, V6T 1Z4, Canada; e-mail: jwerker@psych.ubc.ca

included to highlight the referential nature of the words (Fennell, Waxman & Weisleder, 2007), or if the words are familiarized as parts of longer, clearly distinct words before being taught (Thiessen, 2007). These findings generally implicate task demands and give rise to two specific explanations for infants' failure in the standard task. The first possibility is that the Switch failure results from insufficient learning. Although infants are capable of learning new labels for objects (noticing a swap of *lif* for *neem*; Stager & Werker, 1997), some of the phonetic features of the new words may not be encoded into the memory representations. This lack of encoded detail could result in indistinguishable representations for sufficiently similar word forms. This account pins the difficulties on the learning phase and suggests that aspects of word learning in the infant's natural environment, such as exposures that are more variable or distributed over relatively long periods, provide superior learning to the concentrated exposure provided in the laboratory training task. Familiar words, learned in a natural environment prior to laboratory exposure, would draw on more detailed representations.

Alternatively, it is possible that 14-month-olds can handle the learning phase and successfully encode the phonetic features of each of the words, linking each word more closely to its taught referent than the other word. But the distinction between the words might not be encoded with sufficient confidence to lead to a dishabituation response on the Switch trials. In this case failure would be closely tied to the demands of the Switch test phase; it is the *application* of the variation children encode that is the limiting factor in 14-month-olds' performance. Specifically, the Switch task requires that infants reject an object (e.g. the 'bih') being labelled with a different but very similar-sounding name (*dih*). Yet infants could have learned that *bih* is a better match to the 'bih' object than *dih* without triggering greater looking to the switch; the similar word could be judged to be a plausible instance of the newly learned label. This account suggests that the learning ability of 14-month-old infants is greater than previously attributed and that the test phase could have masked an ability to learn two similar-sounding words in an experimental session. On this account infants would be able to associate new words with referents and encode the phonetic detail of the word form to some degree, but may not consistently respond to mislabelled referents in the task (a possibility discussed in greater detail below).

One way to distinguish between the learning versus test phase accounts for 14-month-olds' failure would be to teach infants with the same Switch learning procedure and change the test phase. The different test phase used in the present study is a two-alternative choice task comparing visual fixations to target and distracter objects. This visual choice task (sometimes called 'Preferential Looking'; Golinkoff, Hirsh-Pasek, Cauley & Gordon, 1987, or 'looking-while-listening'; Fernald, Perfors & Marchman, 2006) has predominantly been used to test phonetic detail in words that are familiar (Mani & Plunkett, 2007; Swingley & Aslin, 2000, 2002). Infants are shown pictures of two objects with phonetically distinct labels side-by-side and the correct label for one of them is heard ('Where's the *dog?*'). A mispronunciation condition presents a label that differs by one sound (*tog*). Looking times are compared across the correct pronunciation (CP) and mispronunciation (MP) conditions. If the children in the CP condition look significantly more to the matching object than children in the MP condition, this indicates that children's stored word representations match the CP better than the MP, and that this difference affects children's comprehension of the word. Whereas the Switch task relies on infants' judgment as to whether a novel object/label pairing is a possible match (necessitating a sufficiently strong 'mismatch' signal to reject the presented pairing on Switch trials), visual choice can tap the degree to which one of the two visible objects is a better match for a presented sound. For example, even if the 'din' object is acceptable as an instance of *bin*, the 'bin' will still be the better match. In studies using visual choice, infants of 14 months detect the mispronunciation of familiar words (Swingley & Aslin, 2002; Swingley, in press). However, as noted above, infants of 14 months also succeed with familiar words in the Switch task (Fennell & Werker, 2003), only failing when taught novel words.

Ballem and Plunkett (2005) have recently used a visual choice paradigm to examine use of phonetic detail in novel words. In their study, 14-month-olds were familiarized to two phonetically distinct word–object pairs (*tuke* and *vope*) and tested on both CPs and MPs. In the first half of the test trials, infants did not evince any recognition of the CPs, whereas in the second half, they did look at the named target picture significantly above chance. This above-chance target fixation response was not found in the MP condition, suggesting that the infants had successfully distinguished between the CP and MP realizations of the target words. However, there was no significant difference between the CP and MP conditions. Coupled with the infants' weak performance overall (absence of target fixation in the CP condition in the first half of test trials), particularly following a learning phase that did not require phonetic disambiguation, the failure to fixate the target when given a mispronunciation in the second half of test trials must be interpreted cautiously.[1]

The current experiment teaches 14-month-olds two similar-sounding words using the standard Switch

[1] Mani and Plunkett (2008) taught 14-month-olds two words (*padge*, *mot*) with four training exposures to each word. Infants were then tested in a visual choice design and recognized the words when pronounced correctly but not when mispronounced as *poudge* and *mit*. This reveals that infants can encode at least some information about a word's vowels even after only a few exposures, though it does not demonstrate that the vowels were necessarily encoded with a high level of accuracy because the deviant vowels tested as MPs were different from the trained vowels in several phonetic dimensions.

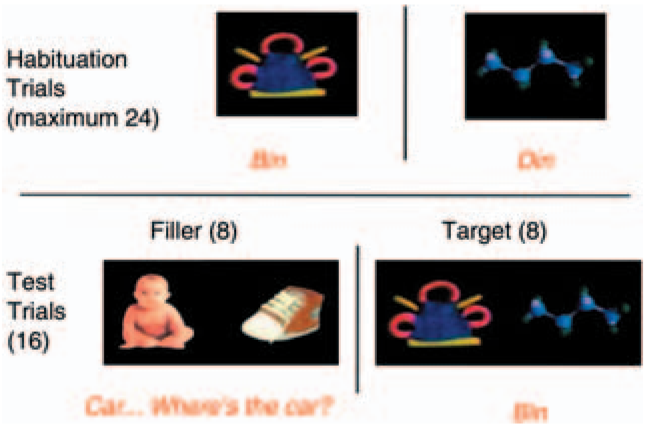

Figure 1 *Design.*

habituation procedure, and tests using a visual choice design. The MacArthur-Bates Communicative Development Inventory (MCDI; Fenson, Dale, Reznick, Bates, Thal & Pethick, 1994) was used to assess infants' prior word learning, to examine whether children with larger vocabularies show larger and more reliable word learning effects – a result previously found using the standard Switch dishabituation task (Werker *et al.*, 2002). Teaching the words using the typical Switch habituation, rather than a familiarization, entails a rough calibration to individual learning rates, possibly increasing the likelihood that infants will learn the words. The use of the visual choice test will reveal whether infants' previous failure at 14 months is specifically related to the Switch test phase, and could yield the first evidence that 14-month-olds can simultaneously learn two similar-sounding words.

Method

Participants

Thirty-six 14-month-old infants completed the study (18 girls, *M*: 14 months 12 days, range: 13 months 29 days–15 months 0 days) in Vancouver, Canada. Each infant was reportedly exposed to English at least 80% of the time, with only eight infants having some exposure to another language. Infants were all full-term and no health problems were reported. An additional 35 infants were excluded from analyses for the following reasons: fussiness ($n = 28$), equipment failure ($n = 6$), and failure to habituate ($n = 1$).

Apparatus

The study was conducted in a 280 × 226 cm sound-attenuated room painted dark mauve and lit by 15 W fluorescent lamps mounted to each side. Images were projected onto a front screen surrounded by black cloth. A digital video camera peeked out of a small hole in the cloth below the screen, relaying the image of the child to the control room and recording looking behaviour. Audio speakers were hidden behind the curtains to the left and right of the projector playing sounds between 68 and 72 dB SPL.

The experiment was run using a Macintosh G4 computer with the Habit 2000 program (Cohen, Atkinson & Chaput, 2000) controlling stimulus presentation and collecting on-line looking time data. The video recordings were later digitized using Final Cut Pro for off-line coding of looking behaviour in a frame-by-frame analysis (30 frames per second).

Procedure

Infants sat on their parents' lap 1.1 metres from the screen. Parents wore Koss TD/65 headphones playing female vocal music. The experimenter was seated in an adjoining control room, unable to hear the auditory stimulus being presented on any given trial.

The experimenter initiated each trial when the infant visually fixated on a flashing neon green light in the centre of the screen. The study began with a pretest comprising a rotating toy waterwheel and the nonsense word *neem*. The habituation phase consisted of a quasi-randomly ordered series of trials, with each trial containing one of two word–object pairs. There were two trials of each word–object pairing in every block of four trials. Both target objects had been previously used in Switch experiments (Werker *et al.*, 2002), and moved slowly up and down in synchrony. The auditory stimuli were two labels unlikely to be familiar to infants, differing only in the initial consonant: *bin* and *din* (Pater *et al.*, 2004). All auditory stimuli were produced in an infant-directed speech register by a female American English speaker. Stimuli were recorded in a soundproof booth onto an Intel-based PC using Kay Elemetrics Computerized Speech Lab (CSL) software at 44 kHz. Each trial contained seven tokens of the label, each spoken in a different intonation, with the first three tokens repeated at the end. Labels were not consistently presented at the moment at which the object changed direction (Gogate & Bahrick, 1998). The habituation criterion was set at a decrease of 65% of that of the longest previous block (Stager & Werker, 1997), with a maximum of 24 trials.

The test trials featured both moving objects on the screen at once, with only one named. On each target test trial one word (*bin* or *din*) was presented in isolation a total of four times. The test phase was supplemented with similarly composed filler trials of familiar objects (*car*, *shoe*, *dog*, and *baby*). The filler tokens were pronounced once in isolation, then in a carrier sentence (e.g. *Car! Look at the* car*!*) to help maintain infants' interest in the procedure and help them orient to the task. There were 16 test trials in all, eight filler and eight target (both counterbalanced for side over the entire testing phase). Two filler trials were always presented first, counterbalanced for side (i.e. the left and the right objects were each named once), the third and fourth trials were target trials, and the rest were randomly ordered (see Figure 1).

Parents completed the MCDI (Fenson *et al.*, 1994) within 1 week of participation as a measure of each infant's receptive and productive vocabularies.

Results

Vocabulary analysis

MCDIs were obtained for 27/36 infants. Reported comprehension vocabulary ranged from 25 to 247 words (M = 102, SD = 62.0) and reported production ranged from 0 to 68 words (M = 15, SD = 17.6).

Test trials

The target test trials and filler trials were analyzed over 367–2000 ms after the onset of the spoken target word (e.g. Swingley & Aslin, 2002). The average proportions of fixation to the target (correct) object were computed separately for the target test and filler trials and compared to chance (50%).

Target test trials

One outlier lay 2.7 standard deviations from the mean and was excluded from all analyses. Over all eight trials, the mean fixation to the correct object was 53.5%, significantly different from chance [$t(34) = 2.31$, $p = .027$, 22/35 infants looking more than 50%]. Next, the first and second blocks were analyzed separately (Ballem & Plunkett, 2005). Infants' fixation to the correct object was above chance in the first four target test trials [$t(34) = 2.79$, $p = .009$; M = 56.8%, 25/35 infants] but not in the second four trials [$t(34) = -.185$, $p = .85$; M = 49.4%; 17/35 infants].[2]

Correlations were explored between looking time data and parents' reports of receptive and productive vocabulary sizes on the CDI. Comprehension and production scores were both positively correlated with looking to the correct object over all target test trials (comprehension $r = .409$, $p = .034$; production $r = .413$, $p = .032$). In a block analysis there was no correlation in the first block of target test trials (comprehension $r = .075$, $p = .71$; production $r = .121$, $p = .55$) but there were stronger correlations over the second block (comprehension $r = .329$, $p = .094$; production $r = .489$, $p = .01$).

[2] Because both objects in any picture pair served both as target and distracter on separate trials, 50% target fixation may be taken as the overall chance level (pictures served as their own controls). For comparability with some previous work, we also compared, trial by trial, the difference between looking during the test window and looking prior to the spoken target word. If children understand the word, this difference should exceed zero. This analysis reveals the same pattern of results, with infants succeeding over all eight trials [$t(34) = 2.604$, $p = .014$], and in the first block, but not the second [$t(34) = 2.116$, $p = .042$ in the first block; $t(34) = -.819$, $p = .42$ in the second block].

Filler trials

Over all filler trials, mean fixation to the correct object in the window was 54.2%, marginally different from chance [$t(34) = 1.833$, $p = .076$; 23/35 infants looked more than 50%]. However, for all infants, a filler trial was the first test trial following habituation; a move from a single object display to a dual object display. If this first trial is excluded as a 'warm-up' trial, fixation to correct becomes 56.4%, significantly different from chance [$t(34) = 2.358$, $p = .024$; 21/35 infants]. Further, the filler trials were intended to present items for which infants had already attached labels. In fact, of the 27 infants for whom MCDIs were obtained, 21 were rated by their caretaker to understand *baby*, 22 knew *car*, 25 knew *dog*, and 21 knew *shoe* (one infant knew none of the four words). If only the data for which infants were rated to understand the word are included, over eight trials, mean fixation is 58.0% to the correct object [$t(25) = 3.33$, $p = .003$, 19/26 infants]; when excluding the first trial, mean fixation is 60.0% to the correct object [$t(25) = 3.10$, $p = .005$, 19/26 infants].

Habituation trials

The mean number of trials to reach the habituation criterion was 12.67 trials ($SD = 4.33$, range: 8–24). Correlations revealed trends with performance on the second block of filler test trials when only data for which infants were rated to have understood the words were included ($r = -.377$, $p = .058$), and with performance on the first block of target test trials ($r = .277$, $p = .101$). However, number of trials to habituate did not correlate with MCDI comprehension or production data, performance on target test trials, or performance on other filler trial data, even when only data for which infants were rated to understand the words were included (all ps > .15).

Discussion

This study taught 14-month-old infants two phonetically similar words for two new objects. Prior research had failed to show evidence of phonetic detail in word learning under these conditions. Here, a visual choice test phase was used instead of the typical Switch test, and the infants succeeded. This success confirms infants' ability to detect phonetic detail in word learning, and suggests that infants' learning ability was previously masked by the demands of the testing phase in the Switch procedure.[3]

[3] In principle, maybe a 360–2000 ms post-word onset analysis period would also reveal significant recovery from habituation in the Switch procedure. However, analyses of the shorter window conducted on the raw 14-month-old data from Werker *et al.* (2002) do not reveal any improvement in performance.

Why do 14-month-old infants fail on the Switch but not the visual choice paradigm? We suggest two possibly complementary hypotheses. One is consistent with descriptions of difficulty in the Switch task as arising from resource demands (Casasola & Cohen, 2000; Fennell & Werker, 2003; Werker *et al.*, 2002). The fact of 14-month-olds' Switch success with familiar words indicates that the differences between 14-month-old and 17-month-old infants do not result from 14-month-olds being utterly unable to encode phonological properties of words under the present teaching conditions. They do not lack the relevant representational resources. Rather, on this account, 14-month-olds are less successful under the demanding learning conditions of the Switch task, and differences in performance could be expected by varying task difficulty. Indeed, phonetic detail of even familiar words is seen to suffer in a difficult task that lacks the referential context of seeing the named objects (Mills, Prat, Stager, Zangl, Neville & Werker, 2004), and while 14-month-olds fail on the Switch task whether one or two words are taught (Stager & Werker, 1997), a variant on the Switch task does result in success on the presumably easier one-object procedure, but not on the two-object version (Fennell *et al.*, 2007).

The present findings show that infants' representations of the newly learned words are distinct from one another. These representations may nevertheless be relatively noisy, or encoded with relatively low confidence, resulting in uncertainty about whether the mismatching word is truly an inappropriate name for the object. Presumably there are also natural conditions under which children come away from a linguistic interaction with only partially complete phonological representations (see Swingley, 2007). A probabilistic account that supposes non-binary feature representations suggests that an infant may be, for example, 60% certain that the 'bin' object was [bilabial] and 60% certain that the 'din' object was [alveolar]. In the Switch task, if presented with the 'bin' object and the *din* label, this 60% mismatch probability, like the 40% uncertainty generated by the matching *bin* label, may not surpass the threshold level required for rejection of the pairing, resulting in no difference between the same and Switch trials. In the case of familiar words, linking the referential and phonetic properties would be a less demanding task under which the infant may reach a higher level of certainty, for example, 75%. This would presumably surpass the threshold for rejection in the mismatch condition, resulting in a difference with the sub-threshold match condition (25%), in contrast to performance with newly learned words. In the visual choice paradigm, the critical criterion involves a direct comparison of, and a forced choice between, the two objects. Here, the relatively slight difference in uncertainty is sufficient; the 60% match would be considered to be a better fit than the 40% match, allowing infants to demonstrate their learning ability.

An alternative to the general cognitive resources explanation is that phonological interpretation *per se* is what limits 14-month-olds' performance. This explanation centres on the distinction between *phonetic* and *phonological* representation: the former being a characterization of words' forms interpreted according to their acoustic properties as perceived by the auditory system, and the latter being a characterization of words as composed of discrete phonological categories interpreted within a framework in which single-category differences signal lexical distinctions. Phonetic learning has roots in a bottom-up analysis of speech sounds in the first year of life (Anderson, Morgan & White, 2003; Maye, Werker & Gerken, 2002; Werker, Pons, Dietrich, Kajikawa, Fais & Amano, 2007), prior to the acquisition of a meaningful vocabulary (Kuhl, 2004; Werker & Tees, 1999), and children encode phonetic details of familiar words before they know minimally different words that exemplify phonological distinctions (Swingley, 2003). This early ability could account for 14-month-olds' ability to detect changes in familiar words; sensitivity to the typical phonetic variability of the word itself could cue infants to an out-of-the-ordinary change as suspect – but only within the context of the familiar word. That is, the younger infants may have accumulated enough experience with familiar words such as *ball* to know that it is never pronounced *doll*, but this specific knowledge has yet to be generalized to all *b–d* contrasts. Thus, the general principle that phonological differences signify different words may not be available to 14-month-olds learning new words like *bin* and *din*, leaving them with the same types of nonbinary probabilistic phonetic representations as hypothesized above. On this view, calling a *din* object by the word *bin* is not good pronunciation to the 14-month-old, but neither is it categorically incorrect.

Over the course of the second year, children become increasingly likely to assume that phonological differences signal different words. Performance on new words after 14 months is transitional, marked by success under some conditions (Curtin, Fennell, Escudero & Werker, in press; Dietrich, Swingley & Werker, 2007; van der Feest, 2007; Werker *et al.*, 2002; White & Morgan, 2008) and failure under others (Swingley & Aslin, 2007). This phonological development may be directed by other sources of information, including the top-down structure of the vocabulary (Swingley & Aslin, 2007; Werker & Curtin, 2005; Yeung & Werker, under review) and the phonological distribution of speech sounds in the lexicon (Thiessen, 2007). Note that this development could also contribute to 17-month-olds' success in a scenario where encoding difficulties result from insufficient general processing ability, as the emerging phonological knowledge could ease the required cognitive resources to reject an incorrect pairing (Werker & Curtin, 2005).

The current study revealed a relationship between vocabulary size (both productive and receptive) and successful performance with the phonetically similar new words. These correlations are consistent with both positions outlined above. Infants with larger vocabularies could have more easily applied phonetic detail in new words because they are better or more experienced word learners, easing the overall cognitive burden and allowing for clearer recognition of the phonetic distinction. Additionally, larger vocabulary may be indicative of greater phonological sophistication, allowing for more efficient application of this knowledge. The presence of a relationship is consistent with previous Switch work where success in differentiating similar new words correlates with age and vocabulary size in infants, both concurrently (Werker *et al.*, 2002) and predictively (Bernhardt, Kemp & Werker, 2007). The relationship between vocabulary and performance seen in the current study is most apparent in the later test trials, consistent with the suggestion that phonological knowledge may help guard against perceptual adaptation (Thiessen, 2007). These benefits of increased vocabulary size appear to be restricted to new words; concurrent vocabulary measures have not been predictive of detection of stimulus mispronunciations of familiar words (Bailey & Plunkett, 2002; Swingley & Aslin, 2000, 2002), although it is clear that familiar word performance in the visual choice task does improve with age and vocabulary (Fernald *et al.*, 2006; Fernald, Pinto, Swingley, Weinberg & McRoberts, 1998).

The use of a visual choice test phase revealed that 14-month-old infants are capable of concurrently learning two phonetically similar words; however, the comparison with Switch procedures cautions that this ability is not strong, and is still developing. Yet the present findings clearly demonstrate that novice word learners can detect and incorporate sufficient phonetic detail into new word forms. We have described two types of explanation for the early failure: one which supposes that transition to success at 17 months results from development in general cognitive abilities, allowing infants to encode auditory and visual information more accurately, and another which attributes success to gradual learning of phonological principles of interpretation, namely the fact that a change from a *b* to a *d* sound always signals a lexical difference. Disentangling these possibilities to understand how the ability to use phonetic detail in word learning develops and strengthens is an important goal of future research.

Acknowledgements

Thanks to all the parents and infants who participated, and to Lila Gleitman, Anne Cutler, Marisa Cruickshank, Laurel Fais, Jason Lau, Clarisa Markel, Erin Moon, and Heidi Ostberg for their essential contributions to this work. This research was supported by Natural Sciences and Engineering Research Council of Canada (NSERC) grant RGPIN1103-05 to JFW, McDonnell Foundation grant 220020096 to JFW, National Institutes of Health grant R01-HD049681 to DS, National Science Foundation grant HSD-0433567 to Delphine Dahan and DS, an NSERC graduate fellowship to CTF, and Social Sciences and Humanities Research Council of Canada, Michael Smith Foundation for Health Research, and British Columbia Medical Services Foundation graduate fellowships to KAY.

References

Anderson, J.L., Morgan, J.L., & White, K.S. (2003). A statistical basis for speech sound discrimination. *Language and Speech*, **46**, 155–182.

Bailey, T., & Plunkett, K. (2002). Phonological specificity in early words. *Cognitive Development*, **17**, 1267–1284.

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

Bernhardt, B.M., Kemp, N., & Werker, J.F. (2007). Early word–object associations and later language development. *First Language*, **27**, 315–328.

Burns, T.C., Yoshida, K.A., Hill, K., & Werker, J.F. (2007). The development of phonetic representation in bilingual and monolingual infants. *Applied Psycholinguistics*, **28**, 455–474.

Casasola, M., & Cohen, L.B. (2000). Infants' association of linguistic labels with causal actions. *Developmental Psychology*, **36**, 155–168.

Cohen, L.B., Atkison, 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.

Curtin, S., Fennell, C.T., Escudero, P., & Werker, J.F. (in press). Weighting of vowel cues explains patterns of word–object associative learning. *Developmental Science*.

Dietrich, C., Swingley, D., & Werker, J.F. (2007). Native language governs interpretation of salient speech sound differences at 18 months. *Proceedings of the National Academy of Sciences*, **104**, 16027–16031.

Fennell, C.T. (2004). Infant attention to phonetic detail in word forms: knowledge and familiarity effects. Unpublished doctoral dissertation, the University of British Columbia, Vancouver.

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., Waxman, S.R., & Weisleder, A. (2007). With referential cues, infants successfully use phonetic detail in word learning. *Proceedings of the 31st Annual Boston University Conference on Language Development*.

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

Fenson, L., Dale, P.S., Reznick, J.S., Bates, E., Thal, D.J., & Pethick, S.J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, **5**, 1–173.

Fernald, A., Perfors, A., & Marchman, V.A. (2006). Picking up speed in understanding: speech processing efficiency and vocabulary growth across the second year. *Developmental Psychology*, **42**, 98–116.

Fernald, A., Pinto, J.P., Swingley, D., Weinberg, A., & McRoberts, G. (1998). Rapid gains in speed of verbal processing by infants in the second year. *Psychological Science*, **9**, 228–231.

Gogate, L.J., & Bahrick, L. (1998). Intersensory redundancy facilitates learning of arbitrary relations between vowel sounds and objects in seven-month-old infants. *Journal of Experimental Child Psychology*, **69**, 133–149.

Golinkoff, R.M., Hirsh-Pasek, K., Cauley, K.M., & Gordon, L. (1987). The eyes have it: lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, **14**, 23–45.

Jusczyk, P.W., & Hohne, E.A. (1997). Infants' memory for spoken words. *Science*, **277**, 1984–1986.

Kuhl, P.K. (2004). Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience*, **5**, 831–843.

Kuhl, P.K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language perception between 6 and 12 months. *Developmental Science*, **9**, F1–F9.

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**, 53–59.

Maye, J., Werker, J.F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, **82**, B101–B111.

Mills, D., Prat, C., Stager, C., Zangl, R., Neville, H., & Werker, J. (2004). Language experience and the organization of brain activity to phonetically similar words: ERP evidence from 14- and 20-month-olds. *Journal of Cognitive Neuroscience*, Special Issue on Developmental Cognitive Neuroscience, Charles A. Nelson (Ed.), **16**, 1452–1464.

Narayan, C. (2006). Acoustic-perceptual salience and developmental speech perception. Unpublished doctoral dissertation.

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

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. (2003). Phonetic detail in the developing lexicon. *Language and Speech*, **46**, 265–294.

Swingley, D. (2005). 11-month-olds' knowledge of how familiar words sound. *Developmental Science*, **8**, 432–443.

Swingley, D. (2007). Lexical exposure and word-form encoding in 1.5-year-olds. *Developmental Psychology*, **43**, 454–464.

Swingley, D. (in press). Onsets and codas in 1.5-year-olds' word recognition. *Journal of Memory and Language*.

Swingley, D., & Aslin, R.N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, **76**, 147–166.

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

Swingley, D., & Aslin, R.N. (2007). Lexical competition in young children's word learning. *Cognitive Psychology*, **54**, 99–132.

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

van der Feest, S.V.H. (2007). Building a phonological lexicon: the acquisition of the Dutch voicing contrast in perception and production. Doctoral Dissertation, Raboud University Nijmegen. Utrecht: Prince Productions BV.

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**, 1289–1309.

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

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

Werker, J.F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L., & Amano, S. (2007). Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition*, **103**, 147–162.

Werker, J.F., & Tees, R.C. (1999). Experiential influences on infant speech processing: toward a new synthesis. In J.T. Spence (Ed.), J.M. Darley & D.J. Foss (Associate Editors), *Annual review of psychology*, **50** (pp. 509–535). Palo Alto, CA: Annual Reviews.

White, K.S., & Morgan, J.L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language*, **59**, 114–132.

Yeung, H.H., & Werker, J.F. (under review). Learning words' sounds before learning how words sound: 9-month-olds use distinct objects as cues to categorize speech information.

Received: 29 September 2007
Accepted: 3 April 2008