

# Native language governs interpretation of salient speech sound differences at 18 months

Christiane Dietrich\*, Daniel Swingley<sup>†‡</sup>, and Janet F. Werker<sup>§</sup>

\*Max-Planck-Institute for Psycholinguistics, Wundtlaan 1, 6500AH, Nijmegen, The Netherlands; <sup>†</sup>Department of Psychology, University of Pennsylvania, 3401 Walnut Street, 302C, Philadelphia, PA 19104; and <sup>§</sup>Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver, BC, Canada V6T 1Z4

Edited by Dale Purves, Duke University Medical Center, Durham, NC, and approved August 13, 2007 (received for review June 7, 2007)

One of the first steps infants take in learning their native language is to discover its set of speech-sound categories. This early development is shown when infants begin to lose the ability to differentiate some of the speech sounds their language does not use, while retaining or improving discrimination of language-relevant sounds. However, this aspect of early phonological tuning is not sufficient for language learning. Children must also discover which of the phonetic cues that are used in their language serve to signal lexical distinctions. Phonetic variation that is readily discriminable to all children may indicate two different words in one language but only one word in another. Here, we provide evidence that the language background of 1.5-year-olds affects their interpretation of phonetic variation in word learning, and we show that young children interpret salient phonetic variation in language-specific ways. Three experiments with a total of 104 children compared Dutch- and English-learning 18-month-olds' responses to novel words varying in vowel duration or vowel quality. Dutch learners interpreted vowel duration as lexically contrastive, but English learners did not, in keeping with properties of Dutch and English. Both groups performed equivalently when differentiating words varying in vowel quality. Thus, at one and a half years, children's phonological knowledge already guides their interpretation of salient phonetic variation. We argue that early phonological learning is not just a matter of maintaining the ability to distinguish language-relevant phonetic cues. Learning also requires phonological interpretation at appropriate levels of linguistic analysis.

language acquisition | lexical development | phonology | word learning

Infants begin to acquire their native language by learning phonetic categories (1, 2). At birth, infants seem to distinguish most of the phonetic contrasts used by the world's languages. However, over the first year, this "universal" capacity shifts to a language-specific pattern in which infants retain or improve categorization of native-language sounds but fail to discriminate many nonnative sounds. Children's failure to discriminate non-native sound contrasts is actually advantageous for learning language, because it prevents children from misinterpreting within-category phonetic variation as indicating a linguistic distinction.

However, whereas infants' phonetic category learning is of clear relevance to language acquisition, discrimination failure is not the only mechanism that is required for developing a native-language phonology that can correctly categorize words (3–5). Many phonetic properties of speech remain discriminable to infants even when the native language does not use these properties to distinguish words. For example, pitch variation is a salient phonetic property of speech but is not used systematically in English for marking different words (in contrast to "tone" languages, like Mandarin, in which syllables comprising the same phonetic segments have wholly different meanings depending on their tone). Vowel duration is another example. Vowel duration is a salient phonetic property that is used to distinguish words in some languages (such as Japanese and Finnish) but not others (such as English and French).<sup>¶</sup> For cases

like pitch and duration variation, the problem to be solved is not about ignoring variation, but assigning it the correct linguistic function. English vowel duration, for example, is used in signaling linguistic stress (6, 7), and assigning prosodic structure (8–10), speech features children must learn to understand and produce (11–13).

The present series of experiments investigated how very young children begin to interpret discriminable phonetic variation in language-specific ways when learning words. We studied 18-month-olds because, at this age, children successfully learn words in the task we used (14) but have relatively small vocabularies (allowing us to evaluate lexicon-based accounts of phonological development, as described in *Discussion*). There are indications that, under some circumstances, infants begin to weight phonologically relevant phonetic variation more heavily than irrelevant variation even before 18 months. For example, 10.5-month-olds, but not 7.5-month-olds, recognize novel instances of recently encountered words even when the novel instances differ in talker affect [e.g., happy vs. neutral; (15)] or talker sex (16). However, 18-month-olds are less constrained than older children in their willingness to interpret nonlinguistic symbols like gestures, pictograms, and even beeps as if they were words (ref. 17; see also ref. 18), suggesting that they might also interpret salient phonetic variation as signaling a lexical distinction even if this interpretation was not licensed by the language's phonology.

The phonetic variable we tested was vowel duration. Participants were native learners of either English or Dutch. In English, duration is not the primary cue to vowel identity, although it is apparent in some contexts as a secondary cue (19, 20). By contrast, in Dutch, duration is an important cue used in differentiating the low vowels [a] and [a:], although it is not the only cue (e.g., see ref. 21).<sup>||</sup> Phonetic studies suggest that differences between the long and short vowels of Dutch are larger than any analogous differences for English (22). Informally, as well, Dutch native speakers consistently affirm that the ad libitum vowel lengthening characteristic of American English infant-directed speech (*Where's the baaaaall!*) is not natural in Dutch, and is excluded for words containing the [a] vowel (presumably because such lengthening would change the vowel's phonological category, or at least render the vowel ambiguous). As a result of

Author contributions: C.D. and D.S. designed research; C.D., D.S., and J.F.W. performed research; J.F.W. contributed new reagents/analytic tools; C.D. and D.S. analyzed data; and D.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

<sup>†</sup>To whom correspondence should be addressed. E-mail: swingley@psych.upenn.edu.

<sup>¶</sup>Vowel duration is a cue to the voicing of some following consonants, e.g., to differentiate *hat* from *had*. But in perception of clear speech, consonantal features outweigh vowel duration, and parents do say *hat* with a very long vowel. Here, we tested words ending in [m], so this issue does not arise. Ongoing work not described here tests other consonant contexts.

<sup>||</sup>The Dutch [a] is similar to the vowel in American English *hot*; the Dutch [a:] has no American English equivalent but is somewhat closer to the English vowel of *hat* than the [a] is.

© 2007 by The National Academy of Sciences of the USA

these phonological differences between the two languages, English and Dutch learners face different tasks in learning their phonology. Dutch learners should interpret vowel duration as lexically contrastive in low vowels; English learners should not, even though at this age they readily discriminate vowels varying in duration (R. Mugitani, F. Pons, C.D., J.F.W., and S. Amano, unpublished data).

Word learning was tested by using the Switch procedure, a habituation method that permits within-subjects comparison of responses to novel and familiar pairings of auditory stimuli with visual stimuli (24). Children were habituated to two scenes: one in which a novel object was named using varied tokens of a particular syllable, and one in which a different novel object was named using varied tokens of a different syllable. Upon habituation, children were presented with four test trials. On two of these trials (the “same” or baseline trials), children viewed the same word–object pairings shown during habituation. On two other test trials (the “switch” trials), children viewed audiovisual scenes in which the word–object pairings were swapped. Prior research shows that, when the two words differ by a native-language sound contrast (as in English “bin” vs. “din”), 18-month-olds reliably gaze longer at the display on switch trials relative to baseline trials, indicating recovery of habituation (e.g., see ref. 14).

Here, the syllable pairs that were tested for contrastive interpretations varied either in their vowel duration only (experiments 1 and 2), or in their vowel quality (experiment 3). Because the task evaluates children’s interpretation of words as names for objects, and not merely their discrimination of speech sounds, it is ideally suited to the present question (25). If 18-month-olds have phonological systems that support interpretation of phonetic variation according to the demands of the native language, we would expect Dutch learners, but not English learners, to treat vowel-duration variants as different words. Alternatively, if children use any salient phonetic variation as indicating lexical contrast, learners of both languages should interpret vowel duration as signaling two words.

Three experiments were conducted, each with separate samples of Dutch-learning and English-learning children. In each experiment, children were habituated to two alternating audiovisual scenes: one moving object with repeated presentation of a word, and a second moving object with repeated presentation of another word. The object films were the same in all studies. Auditory stimuli are described in more detail in *Methods*.

Once children habituated, the test phase began, including two “switch” trials and two baseline trials. Only children considering a given phonetic alteration to be relevant to word identity were expected to differentiate the two trial types, looking longer on switch trials than baseline trials.

## Results

**Experiment 1. Dutch-language stimuli.** Thirty-six Dutch and English learners’ responses to changes in vowel duration were tested. The auditory stimuli were several different tokens of the word *tam* (phonetically [tɑm]), read by a native Dutch speaker, forming the short-vowel words, and the same tokens with their vowels digitally lengthened by a factor of 1.9 to form the long-vowel words. The word *tam* is a real Dutch word (meaning *tame*) but is very unlikely to be known by 18-month-olds. The 1.9 ratio of short to long durations is comparable with that found in phonetic studies of naturally produced utterance-final [ɑ] and [ɑ:] sounds in Dutch (22, 26). The tested vowel durations were also well within the range typical of these vowels in spontaneously produced child-directed Dutch.

**Table 1. Children’s mean and (standard deviation) looking times, in seconds, for the Switch and Baseline trials in each of the three experiments**

Exp	Contrast	Child’s language	Switch		Baseline	
			Mean	SD	Mean	SD
1	Dutch duration	Dutch	9.23	3.55	5.04	3.14
1	Dutch duration	English	6.66	3.09	7.15	2.36
2	English duration	Dutch	8.16	3.49	5.92	2.94
2	English duration	English	7.34	3.97	8.04	4.29
3	Dutch quality	Dutch	5.72	2.59	4.08	1.98
3	English quality	English	9.31	3.78	6.31	2.99

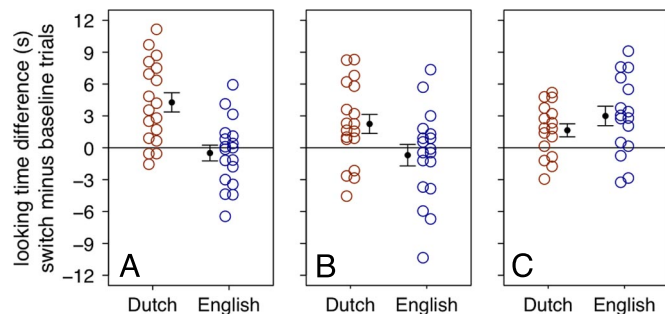
**Findings.** Of the 36 children, 35 (18 Dutch, 17 English) habituated to the audiovisual displays within the maximum 24 trials. \*\* The two groups of children were similar in their total looking times over habituation and test trials. However, only the Dutch children looked at the screen longer on switch trials than on baseline trials. A repeated measures analysis of variance (ANOVA) comparing looking times for the two language groups and trial types revealed a significant interaction between trial type and native language (repeated measures ANOVA  $F_{(1,34)} = 15.71$ ;  $P < 0.0005$ ), as well as a main effect of trial type carried by the Dutch group ( $F_{(1,34)} = 9.77$ ;  $P < 0.005$ ). Mean looking times are given in Table 1. Each child’s looking-time difference score (switch minus baseline trials) is plotted in Fig. 1A. Among the Dutch learners, 15 of 18 (83%) had positive difference scores (binomial  $P < 0.005$ ), with a mean difference score of 4.2 s (SD 3.9); among English learners, 9 of 18 did (50%; mean  $-0.49$ ; SD 3.2). For the Dutch, the two stimulus words yielded similar effects: switch trials presenting the word with the long vowel led to just as much increased attention as the switch trials presenting the word with the short vowel [paired  $t = 0.69$ ; not significant (ns)]. No significant asymmetries were observed in the English sample (paired  $t = 1.34$ ;  $P > 0.15$ ), where neither lengthening nor shortening elicited reliable increases in looking.

Thus, by 18 months, Dutch and English learners treated vowel duration differently in word learning. Dutch learners kept track of the linkage between a given object and a long-voweled syllable, and another object and a short-voweled but otherwise identical syllable. No such effects were found in the English learners, suggesting the influence of the native-language phonology.

**Experiment 2. English-language stimuli.** In experiment 1, only the Dutch listeners differentiated words varying in vowel duration. However, before attributing this effect to Dutch and English learners’ knowledge of their language’s phonology, it is necessary to exclude the possibility that the difference in the two groups came about because only the Dutch were listening to words originally pronounced in their native language. It is conceivable that the English learners detected the foreign nature of the speech sounds and, for example, devoted less attention to them as a result.

Experiment 2 therefore tested similar samples of Dutch and English 18-month-olds by using English-language stimuli. A female native speaker of Canadian English was recorded saying the word *tam* (phonetically [tæm]; rhymes with *cam*) to form the short-vowel words. The vowels in these tokens were elongated as in experiment 1 to form the long-vowel words. The [æ] vowel is not found in Dutch; adult Dutch learners of English tend to interpret it as [ɛ] (as in “bed”). The [æ] vowel also tends to be longer than the Dutch [ɑ]. The longer mean duration is partly

\*\*In this and the other experiments, the statistical significance of the results was unchanged in analyses considering only the habituators.



**Fig. 1.** Distribution of looking time difference scores (switch trials minus baseline trials, in seconds) for all children in each experiment. Open circles show each child's difference score; filled circles with error bars indicate the mean and standard error of the mean for each condition. Dutch children's results are plotted on the left in each panel; English learners' results are plotted on the right. (A) Results of experiment 1, comparing Dutch and English children's responses to vowel duration in Dutch stimuli. (B) Results of experiment 2, showing the same for English stimuli. (C) Results of experiment 3, showing responses to vowel quality changes in native-language stimuli.

due to the fact that English does not contrast vowels by duration: whereas the Dutch [a] is kept short by the contrasting [a:], there is no such constraint on English [æ]. As a result, the absolute size of the duration manipulation, although not the ratio, was greater in experiment 2 than experiment 1 (274 ms vs. 134 ms), and whereas the short vowel was well within the typical child-directed range for an utterance-final [æ], the long vowel was unusually long (although it did not sound unnatural). Thus, the manipulation was not subtle, and therefore provided a strong challenge to the hypothesis suggested by experiment 1: that English learners do not consider vowel duration lexically contrastive.

**Findings.** Of the 36 children in the sample, 30 (17 Dutch, 13 English) habituated to the displays within 24 trials. Nonhabitua-tors' performance matched the modal pattern.<sup>††</sup> As in experiment 1, only the Dutch children fixated the screen longer on switch trials than baseline trials. An ANOVA comparing looking times revealed a significant interaction between trial type (switch, baseline) and native language (repeated measures ANOVA  $F_{(1,34)} = 4.74$ ;  $P = 0.037$ ). No other effects were significant. Mean looking times are given in Table 1, and each child's looking-time difference score is plotted in Fig. 1B. Of Dutch learners, 14 of 18 had positive difference scores (78%; binomial  $P < 0.016$ ) with a mean difference score of 2.2 s (SD 3.8). Of English learners, only 8 of 18 had positive difference scores (44%; ns). As in experiment 1, no asymmetries in the effects of presenting the longer word or the shorter word were observed for either language group.

The results therefore replicate the findings of experiment 1, using a new sample of English and Dutch learners and an English stimulus set. Because the tested vowels and their absolute durations were different from those examined in the first study, experiment 2 shows that the English learners' failure to interpret vowel duration contrastively cannot be attributed to the stimuli having been of foreign origin. Further, Dutch learners' contrastive interpretation is not restricted to just the Dutch [a–a:] distinction, nor to a narrow range of absolute duration values.

**Experiment 3. Native vowel quality contrasts.** In experiments 1 and 2, Dutch-learning but not English-learning toddlers showed evidence of learning associations between objects and words where the words differed only in the durations of their vowels. It is

conceivable that Dutch toddlers are more sensitive to vocalic variation of any sort in the present task. It is also logically possible that differences in the implementation of the experimental procedures, rather than true linguistic differences among English and Dutch learners, were responsible for the cross-language differences in effects. To test these possibilities, experiment 3 used the same methods to test toddlers' interpretation of native-language vowels that varied in their quality rather than just their duration. Thus, all listeners were tested on native-language materials. The *tam* tokens of the first two experiments were used, along with new tokens of the word *tem* (phonetically [tɛm]; rhymes with *them*), recorded by the same talkers. Because both groups of learners were tested on native vowel contrasts, we expected equivalent performance, with both groups detecting inversion of the object labels.

**Findings.** Of the 32 children, 27 (16 Dutch, 11 English) habituated to the displays within 24 trials. Habitua-tors and nonhabitua-tors performed similarly, and the statistical results were equivalent considering only the habitua-tors. Contrary to experiments 1 and 2, in experiment 3 children from both language groups fixated the screen longer on switch trials than on baseline trials. An ANOVA comparing looking times revealed a significant main effect of trial type (switch, baseline): repeated measures  $F_{(1,30)} = 17.7$ ;  $P < 0.0005$ . There was also a significant main effect of language group, with the English group looking longer overall ( $F_{(1,30)} = 11.2$ ;  $P < 0.005$ ). The interaction between trial type and language group was not significant ( $F_{(1,30)} = 1.5$ ; ns). Mean looking times are given in Table 1, and condition difference scores are plotted in Fig. 1C. Of Dutch learners, 12 of 16 had positive difference scores (75%; binomial  $P < 0.05$ ); of English learners, 13 of 16 did (81%; binomial  $P = 0.01$ ). As in the preceding experiments, there were no stimulus-word asymmetries: calling the *tam* object "tem" and calling the *tem* "tam" produced equivalent increases in looking.

The longer looking of the English participants relative to the Dutch ones was not anticipated. We speculate that this result was a cultural effect of greater exposure to television in the daily lives of the English learners. This effect may not have been evident in the first two experiments because there the two word types were interpreted as the same by the English learners, making the procedure less interesting to them and decreasing their overall looking time. Evaluating this speculation lies outside the scope of this project.

To confirm the contrast between English learners' success in differentiating words by their vowel quality and their failure for vowel duration, English learners' switch–baseline difference scores were examined in a one-way ANOVA with one level for each experiment. This comparison yielded a significant effect of experiment ( $F_{(2,49)} = 5.1$ ;  $P < 0.01$ ). This effect arose because English learners' responses to changes in quality differed from their (non)responses to changes in duration (two-sample *t* test, Dutch vowel duration vs. vowel quality,  $t(32) = 3.0$ ;  $P = 0.005$ ; English vowel duration vs. vowel quality,  $t(32) = 2.7$ ;  $P = 0.012$ ; English duration vs. Dutch duration,  $t < 1$ ; ns). An analogous analysis of the Dutch children's difference scores yielded a marginal effect of experiment ( $F_{(2,49)} = 2.8$ ;  $P = 0.073$ ); follow-up *t* tests showed that effects of vowel duration changes were greater than effects of the vowel quality change, at least for the Dutch duration stimuli ( $t(34) = 2.3$ ;  $P < 0.02$ ; other comparisons ns).

Thus, in experiment 3, both English and Dutch toddlers readily learned associations between two novel words and their object referents, even though the novel words differed only in their vowel quality. These results rule out trivial explanations of the two groups' different treatment of vowel duration, including variation in laboratory settings and in overall attention to the task.

<sup>††</sup>Among English learners, neither habitua-tors nor nonhabitua-tors showed any sign of longer looking on switch trials than baseline trials. The habitua-tors' mean (switch minus same) difference score was  $-0.48$  (SD 4.74), with 6 of 13 in the positive direction.



## Discussion

Children in each of the experiments were introduced to two novel words under tightly constrained conditions: all tokens were uttered by the same talker and occurred as one-word utterances, and, during habituation, each consistently appeared with one object. In the first two experiments, vowel duration among the instances of each word type varied little, whereas duration between the word types varied by a factor of 1.9. Thus, children were provided with multiple sources of information suggesting that the words might be distinct. Yet only the Dutch, and not the English, seemed to keep track of the link between each word and its referent. The English learners did not lack the cognitive power to succeed in the task, as shown by experiment 3 and by previous studies testing 18-month-olds on words differing in their initial consonant (e.g., see ref. 14). Rather, the difference between the learners of the two languages arose from phonological generalizations they had made based upon their prior language experience.

At present, it is not clear which aspects of the children's experience gave rise to their different treatment of duration in these experiments. One frequently raised hypotheses about phonological category learning is that it is driven by contrast in the vocabulary. Dutch children might learn that [ɑ] and [a:] are different because (for example) the words [stat] ("city") and [start] ("stands") mean different things and are thus used in different contexts. This hypothesis would be supported by evidence that children know such "minimal pair" words by 18 months. However, children that young do not seem to know many word pairs that could clearly indicate a distinction between [ɑ] and [a:]. The Dutch version of the MacArthur–Bates Communicative Development Inventory (CDI) (27, 28), which contains >500 of children's most frequently learned early words, contains no [ɑ–a:] minimal pairs at all. Analysis of 10 children in the Levelt–Fikkert database of Dutch toddlers' speech (29, 30) revealed only one child <21 months of age who was recorded saying any two words whose canonical forms contrasted only in these vowels, and both of these word tokens, which occurred several weeks apart, were coded as imitations rather than spontaneous speech. Finally, a 25,000-word corpus of Dutch infant-directed speech (31) contains only two contrasting pairs: one pair consisting of an adult's nickname and the function word [a:n] (a preposition and part of a construction indicating the present progressive); and the other pair consisting of the words [slɑp] ("bib") and [sla:p] ("sleep"). Whereas each of these three sources represents only a sample of the likely vocabulary knowledge of the typical Dutch 18-month-old, it seems reasonable to question whether the very consistent contrastive interpretation revealed by the Dutch children in the present experiments could rest on so faint a foundation.

The other current hypothesis is that children begin to induce phonological categories "bottom-up," based on their discovery of clusters of speech sounds in phonetic space using perceptual category induction mechanisms (e.g., see ref. 32). Some kind of distributional learning mechanism is undoubtedly implicated in infants' early phonetic category learning, which begins before infants know enough words for the vocabulary-based hypothesis to be feasible (33–35); see refs. 1 and 36 for proposed learning mechanisms). Note, though, that this language-specific perceptual tuning does not necessarily yield categories that are interpreted phonologically; it is a matter of some debate how the phonetic categories that are learned bottom-up then enter into the child's developing system of linguistic contrasts for use in distinguishing words (e.g., see ref. 37). A necessary condition for such learning to be the driving force behind Dutch children's phonological interpretation in the present studies is that long and short vowels be more clearly separable in Dutch than in English. A preliminary examination of this problem using corpora of

Dutch child-directed speech indicated that the set of long and short instances formed largely overlapping distributions.<sup>‡‡</sup> The phonologically long vowel tends to be acoustically longer than the phonologically short vowel, but this tendency is masked by considerable variation in both categories. Given the lack of clear separation in the Dutch long and short vowels' duration distributions, children probably bring additional information to bear in determining that the [ɑ] and [a:] differ, including spectral cues to the distinction.

The literature on the early development of speech perception has focused primarily on early changes in the discriminability of native and nonnative phonetic contrasts. The severe difficulties that children and adults have in distinguishing some nonnative speech sounds are remarkable and have clear implications for language learners and for our understanding of the human capacity for language (38, 39). But distinguishing sounds is only a small part of the problem language learners face. The core problem is to interpret the linguistic significance of phonetic variation, much of which can be detected without difficulty (5). In acquiring English, for example, learners do not lose the ability to discriminate released and unreleased final stops, upward and downward pitch sweeps, or long and short vowels, even though under most circumstances none of these phonetic variations is a primary cue for differentiating words. In some cases, these features are part of the language's prosodic structure and serve to mark phrase boundaries or aspects of discourse; in some other cases, such features have no clear communicative significance. The task of the child is to correctly attribute this variation to the correct level of linguistic interpretation, namely the one that reveals the talker's intentions as he or she implements the phonological rules of the language. The present experiments show that by 18 months Dutch and English children have already begun to solve a small part of this problem.

## Materials and Methods

**Participants.** Children ranged in age from 17;0 (17 months, 0 days) to 18;27. Dutch children were recruited and tested in Nijmegen, the Netherlands; English learners were recruited and tested in Vancouver (experiments 1 and 3) or Philadelphia (experiment 2). All children were being raised monolingually, hearing at least 90% English or Dutch by parental report. Experiments 1 and 2 each tested 18 Dutch and 18 English learners; experiment 3 tested 16 Dutch and 16 English learners. Mean ages in each experiment were: experiment 1, 550 days (SD 14); experiment 2, 539 days (SD 15); experiment 3, 548 days (SD 16). Over each of the three experiments, the average difference in age between the Dutch and English children was 4 days (ns).

**Stimulus Materials.** The visual stimuli consisted of brief digitized animations of brightly colored objects on a black background: a model molecule made of plastic balls; an abstract shape made of colored modeling clay; and a toy waterwheel. The molecule and clay shape moved up and down at a slow speed; the waterwheel turned.

Auditory stimuli in each experiment consisted of three 20-s-long digitized soundfiles each containing 14 tokens of a monosyllabic word, with each token separated from the next by a silent pause of ≈1 s. All words were recorded by using slow, clear speech and varied intonation. One of the three soundfiles contained the word *neem* (phonetically [nim]; rhymes with *team*). The other two each contained one of two monosyllabic words of the form [t, vowel, m], where the identity of the vowels depended on the experiment.

In experiment 1, auditory stimuli were derived from five tokens

<sup>‡‡</sup>Swingley D, Fifteenth Biennial International Conference on Infant Studies, June 20–22, 2006, Kyoto Japan.

of the nonce word *tam* (phonetically [tɑm]) recorded by a female native speaker of Dutch. The mean duration of these syllables was 425 ms (SD 45). The vowel duration manipulation was implemented by lengthening the vowel of these tokens by a factor of  $\approx 1.9$  (range 1.83 to 1.96) by using the PSOLA algorithm available in the waveform editor Praat (40). This change preserves pitch and vowel quality characteristics while resulting in natural-sounding syllables. The mean short vowel duration was 154 ms (SD 13), and the mean long vowel duration was 288 ms (SD 22).

In experiment 2, auditory stimuli were derived from five tokens of the nonword *tam* (phonetically [tæm], recorded by a female native speaker of Canadian English. The mean duration of the recorded tokens (the short-vowel stimuli) was 624 ms (SD 19). Long vowels were produced by lengthening the short vowels by a factor of 1.9 (range 1.82 to 2.07). The mean short vowel duration was 302 ms (SD 30) and the mean long vowel duration was 576 ms (SD 76).

In experiment 3, the unmanipulated *tam* stimuli of the prior experiments were used, with Dutch children hearing the tokens from experiment 1, and the English learners hearing the tokens from experiment 2. An additional set of *tem* stimuli was recorded by the same talkers. The mean duration of the Dutch *tem* tokens was 419.6 ms (SD 50.7), with vowels averaging 143.7 ms (SD 8.7); for the English learners, 553.5 ms (SD 32.9), with the vowels averaging 202.9 ms (SD 23.7).

All testing was done under similar conditions at the Dutch and English sites. The child and parent were led into a quiet, dimly lit room. The child was seated on the parent's lap facing a large video monitor enclosed by a plain black cloth background. Images on the monitor took up  $\approx 13^\circ$  of visual angle at the infant's distance of  $\approx 1$  m. Speech materials were presented from speakers located just above the video monitor (Vancouver) or below it (Nijmegen, Philadelphia). Children's faces were video-

recorded by using a camera concealed beneath the monitor. Parents and experimenters wore headphones playing masking sounds that prevented identification of the test stimuli. Stimulus presentation was controlled by using the *Habit* software package (23). Each trial was preceded by an animation of a rotating blue and white flower-like shape that played until infants attended to it on the screen; then the trial was started. Trials were 20 s long. The first trial was a warm-up trial on which children saw a brief animation of a toy waterwheel while hearing multiple tokens of the nonce word *neem*. On subsequent habituation trials, children were shown two word-object pairings: one object with one of the test words, or the other object with the other test word. These two types of trial alternated in sequence. For the purpose of determining when children habituated, looking on each trial was measured on-line by an observer's keypress. Once average looking over four trials decreased by 50%, children were considered habituated and the test phase began. In the test phase, children viewed all four combinations of words and objects, two familiar ("baseline") and two swapped ("switch"). After the test phase, a fifth, screening trial presented the waterwheel and the word *neem*, a very distinctive stimulus to which dishabituation was expected in all children still attending to the procedure. No children were removed for insufficient attention to this stimulus.

We thank our laboratory managers Sara Clopton, Dr. Laurel Fais, Angela Khadar, and Jane Park, and the many assistants who contributed to this work, and we are very grateful to the parents and children who made this research possible. Portions of this work formed part of C.D.'s PhD thesis, supervised by D.S. This work was funded by the Max-Planck-Gesellschaft, the Nederlandse Organisatie voor Wetenschappelijk Onderzoek's Spinoza Prize (to Dr. Anne Cutler), National Science Foundation Grant HSD-0433567 (to Dr. Delphine Dahan and D.S.), National Institutes of Health Grant R01-HD049681 (to D.S.), and Canadian Natural Sciences and Engineering Research Council Grant 81103 (to J.F.W.).

- Kuhl PK (2004) *Nat Rev Neurosci* 5:831–843.
- Werker JF, Yeung HH (2005) *Trends Cognit Sci* 9:519–527.
- Kazanina N, Phillips C, Idsardi W (2006) *Proc Natl Acad Sci USA* 103:11381–11386.
- Kuhl PK, Conboy BT, Coffey-Corina S, Padden D, Rivera-Gaxiola M, Nelson T (September 10, 2007) *Philos Trans R Soc London B*, 10.1098/rstb.2007.2154.
- Werker JF, Curtin S (2005) *Lang Learn Dev* 1:197–234.
- Redford MA, Davis BL, Miikkulainen R (2004) *Infant Behav Dev* 27:477–498.
- Umeda N (1975) *J Acoust Soc Am* 58:434–445.
- Klatt DH (1976) *J Acoust Soc Am* 59:1208–1221.
- Klatt DH (1987) *J Acoust Soc Am* 82:737–793.
- Wightman CW, Shattuck-Hufnagel S, Ostendorf M, Price PJ (1992) *J Acoust Soc Am* 91:1707–1717.
- Patel R, Grigos MI (2006) *Speech Commun* 48:1308–1318.
- Pollock KE, Brammer DM, Hageman CF (1993) *J Phonetics* 21:183–203.
- Snow D (1994) *J Child Lang* 24:35–56.
- Werker JF, Fennell CT, Corcoran KM, Stager CL (2002) *Infancy* 3:1–30.
- Singh L, Morgan JL, White KS (2004) *J Memory Lang* 51:173–189.
- Houston DM, Jusczyk PW (2000) *J Exp Psychol Hum Percept Perform* 26:1570–1582.
- Namy LL (2001) *Infancy* 2:73–86.
- Woodward AL, Hoynes KL (1999) *Child Dev* 70:65–77.
- Hillenbrand JM, Clark MJ, Houde RA (2000) *J Acoust Soc Am* 108:3013–3022.
- Werker JF, Pons F, Dietrich C, Kajikawa S, Fais L, Amano S (2007) *Cognition* 103:147–162.
- Nooteboom SG, Doodeman GJN (1980) *J Acoust Soc Am* 67:276–287.
- Adank P, van Hout R, Smits R (2004) *J Acoust Soc Am* 116:1729–1738.
- Cohen LB, Atkinson DJ, Chaput HH (2000) *Habit 2000: A New Program for Testing Infant Perception and Cognition* (Univ of Texas, Austin, TX), Version 2.25c.
- Werker JF, Cohen LB, Lloyd VL, Casasola M, Stager C (1998) *Dev Psychol* 34:1289–1309.
- Stager CL, Werker JF (1997) *Nature* 388:381–382.
- Rietveld T, Kerkhoff J, Gussenhoven C (2004) *J Phonetics* 32:349–371.
- Fenson L, Dale PS, Reznick JS, Bates E, Thal DJ, Pethick SJ (1994) *Monographs of the Society for Research in Child Development* 59(5), No 242.
- Swingle D (2003) *Lang Speech* 46:265–294.
- Fikkert P (1994) *On the Acquisition of Prosodic Structure* (Holland Institute of Generative Linguistics, Leiden, The Netherlands).
- Levelt C (1994) *On the Acquisition of Place* (Holland Institute of Generative Linguistics, Leiden, The Netherlands).
- van de Weijer J (1998) PhD thesis (Katholieke Universiteit Nijmegen, Nijmegen, The Netherlands) (MPI Series in Psycholinguistics 9).
- Younger BA (1985) *Child Dev* 56:1574–1583.
- Bosch L, Sebastián-Gallés N (2003) *Lang Speech* 46:217–243.
- Kuhl PK, Williams KA, Lacerda F, Stevens KN, Lindblom B (1992) *Science* 255:606–608.
- Polka L, Werker JF (1994) *J Exp Psychol Hum Percept Perform* 20:421–435.
- Maye J, Gerken LA, Werker JF (2002) *Cognition* 82:B101–B111.
- Dresher BE (2003) *On the Acquisition of Phonological Contrasts*, eds van Kampen NJ, Baauw S (Landelijke Onderzoeksschool Taalwetenschap, Utrecht, The Netherlands), Vol 1, pp 27–46.
- Doupe A, Kuhl PK (1999) *Annu Rev Neurosci* 22:567–631.
- Werker JF, Tees RC (2005) *Dev Psychobiol* 46:233–251.
- Boersma P, Weenink D (2001) *Glott Int* 5:341–345.