

# Sensorimotor influences on speech perception in infancy

Alison G. Bruderer<sup>a,b,1</sup>, D. Kyle Danielson<sup>b</sup>, Padmapriya Kandhadai<sup>b</sup>, and Janet F. Werker<sup>b</sup>

<sup>a</sup>School of Audiology and Speech Sciences, Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada V6T1Z3; and <sup>b</sup>Department of Psychology, University of British Columbia, Vancouver, BC, Canada V6T1Z4

Edited by Patricia K. Kuhl, University of Washington, Seattle, WA, and approved September 10, 2015 (received for review May 2, 2015)

The influence of speech production on speech perception is well established in adults. However, because adults have a long history of both perceiving and producing speech, the extent to which the perception–production linkage is due to experience is unknown. We addressed this issue by asking whether articulatory configurations can influence infants' speech perception performance. To eliminate influences from specific linguistic experience, we studied preverbal, 6-mo-old infants and tested the discrimination of a nonnative, and hence never-before-experienced, speech sound distinction. In three experimental studies, we used teething toys to control the position and movement of the tongue tip while the infants listened to the speech sounds. Using ultrasound imaging technology, we verified that the teething toys consistently and effectively constrained the movement and positioning of infants' tongues. With a looking-time procedure, we found that temporarily restraining infants' articulators impeded their discrimination of a nonnative consonant contrast but only when the relevant articulator was selectively restrained to prevent the movements associated with producing those sounds. Our results provide striking evidence that even before infants speak their first words and without specific listening experience, sensorimotor information from the articulators influences speech perception. These results transform theories of speech perception by suggesting that even at the initial stages of development, oral–motor movements influence speech sound discrimination. Moreover, an experimentally induced “impairment” in articulator movement can compromise speech perception performance, raising the question of whether long-term oral–motor impairments may impact perceptual development.

language acquisition | perception–production | infancy

The acquisition of language, arguably our most defining human capacity, relies on the seamless exchange of information between production and perception. In their seminal work, Eimas et al. (1) found that from 1 mo of age, human infants are equipped with perceptual sensitivities that enable them to discriminate speech sounds according to the boundaries used in human languages (see refs. 2, 3 for reviews of subsequent work). Within the first year, infant speech perception sensitivities adapt to the ambient language: The process of perceptual narrowing results in a decline in discrimination of nonnative distinctions (4, 5) and an improvement in the discrimination of native speech sound contrasts (6, 7). A similar trajectory is seen for audiovisual speech perception: Very young infants can match heard and seen speech (8–10), but by 9–10 mo, they do so reliably only for native speech sounds (11). Development of speech production progresses similarly. Although the infant vocal tract is anatomically immature and lacks the neuromuscular control of the adult vocal tract (12), the ability to produce communicative sounds (cries) is evident at birth (13, 14) and already reflects characteristics of the language experienced in the womb (15). Across the first year of life, the vocal tract grows in size, and articulatory precision increases (13, 14), while infants' babbling behavior begins to reflect the ambient language (16). Although much is known about the perceptual and productive capabilities of human infants, the extent to which the perceptual and productive systems

inform one another during the prelinguistic period remains largely unknown.

In adults, there is considerable evidence of a bidirectional relationship between the speech production and speech perception systems. Listening to speech activates motor and premotor areas in the cortex (17, 18) and influences speech production in perceptual training (19) and auditory feedback tasks (20). Moreover, the speech production system also impacts speech perception: The shape and movement of an adult's articulators affects phonetic perception, as shown with both behavioral (21, 22) and transcranial magnetic stimulation (TMS) procedures (23, 24). In infants, it has been most commonly assumed that any influence is unidirectional, in that the developing production system relies on and reflects incoming auditory perceptual input (16, 25, 26) (see also ref. 27 for self-organizing systems). The only available infant evidence suggesting an influence from production to perception is correlational: Infants show differential attention to those speech sounds that are in their individual babbling and/or productive repertoires (28, 29). Here we test empirically whether motor processes exert a direct influence on the auditory percept.

The auditory–articulatory cortical connections for processing and producing speech are present from early in life. A structural, dorsal pathway connecting the temporal cortex with the premotor cortex, similar to a dorsal pathway found in adults (30–32), is present in newborns to “support auditory-to-motor mappings” (33). These cortical areas are active during speech perception not only in preverbal infants (34, 35) but also in neonates born as early as 29 wk gestation (36). With the anatomical structure in place to support auditory-to-motor mapping during language acquisition (33), a bidirectional influence could be possible: Exposure to auditory speech influences the articulatory–motor

## Significance

Theories of language acquisition have typically assumed infants' early perceptual capabilities influence the development of speech production. Here we show that the sensorimotor (production) system can also influence speech perception: Before infants are able to speak, their articulatory configurations affect the way they perceive speech, suggesting that the speech production system shapes speech perception from early in life. These findings implicate oral–motor movements as more significant to speech perception development and language acquisition than current theories would assume and point to the need for more research on the impact that restricted oral–motor movements may have on the development of speech and language, both in clinical populations and in typically developing infants.

Author contributions: A.G.B., D.K.D., P.K., and J.F.W. designed research; A.G.B. and D.K.D. performed research; A.G.B., D.K.D., P.K., and J.F.W. analyzed data; and A.G.B., D.K.D., P.K., and J.F.W. 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>1</sup>To whom correspondence should be addressed. Email: agreuel@audiospeech.ubc.ca.

This article contains supporting information online at [www.pnas.org/lookup/suppl/doi:10.1073/pnas.1508631112/-DCSupplemental](http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.1508631112/-DCSupplemental).

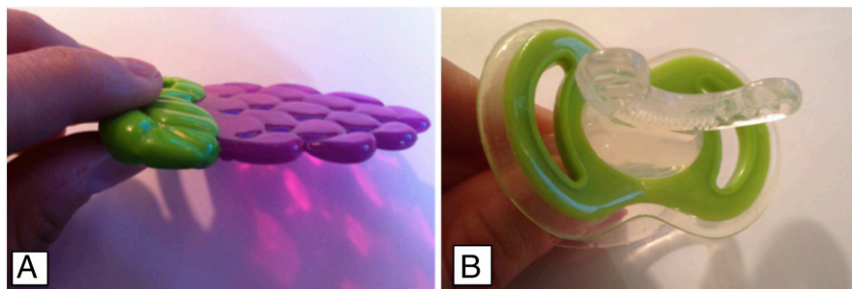


Fig. 1. Teether images. (A) Flat teether. (B) Gummy teether.

linkage, and in turn, sensorimotor information from the articulators could impact the auditory speech percept. Recent work has shown that experience listening to speech reorganizes the way in which motor areas are recruited during infant speech perception, likely built through feedback from hearing the output of one's own nascent productions (37). Although such work reveals an early-appearing link between perception and production, the data support, and are interpreted to indicate, the importance of perception in establishing that link.

The present work provides the first empirical test, to our knowledge, that the influence can also work in the other direction: That speech perception sensitivities in young infants rely in part on the productive system (through specific articulatory configurations). To test this hypothesis, we experimentally manipulated the position of 6-mo-old English-learning infants' articulators and tested whether this affected their discrimination of a speech sound contrast that is nonnative to English. The contrast used was the Hindi dental /d/ versus retroflex /ɖ/ distinction. It has been well established that, although English-speaking adults and English-learning infants older than 10 mo cannot discriminate this distinction, English-learning infants aged 6 mo, who have not yet undergone perceptual narrowing to the native language, can discriminate the distinction (4).

The choice for the Hindi /d/-/ɖ/ contrast in experiments 1–3 is crucial for two reasons: First, this distinction is not used in English; hence, English-learning infants are not exposed to this phonetic contrast and thus have had no specific linguistic experience with /d/ versus /ɖ/. That they have not previously heard these phones controls for the possibility that some earlier occurring perceptual experience primed the link to production. Second, the main articulatory difference between these two sounds is in the placement of the tongue tip during pronunciation. The dental /d/ is produced by placing the tongue at the back of the teeth on the roof of the mouth, and the front of the hard palate, thus in front of the English alveolar /d/. The retroflex /ɖ/ is a subapical consonant, produced by curling the tongue tip back behind the English alveolar /d/, so that the bottom side of the tongue tip makes contact with the roof of the mouth. Therefore, if sensorimotor information shapes infant speech perception, selectively impairing tongue tip movement should also impair auditory discrimination of the Hindi /d/-/ɖ/ contrast.

To temporarily impair the movement and control the placement of infants' tongues, we built on the design of Yeung and Werker (38), in which teething toys were used to render specific mouth shapes (either rounded-lip shapes via pacifiers or spread-lip shapes via wide ring teethers) during an audiovisual speech-matching task in 4-mo-old infants. The use of teething toys to experimentally constrain the shape and movements of the mouth allows for the investigation of prearticulatory sensorimotor influences on speech perception in infancy. Experiment 1 was conducted without any oral-motor manipulation to verify the successful discrimination of the particular auditory recordings of the two Hindi sounds by English-learning infants. To test the specificity of a sensorimotor influence on auditory-only speech perception, a teether that selectively blocked tongue tip movement was used in experiment 2 (Fig. 1A), and a teether that did not interfere with tongue tip

movement was used in experiment 3 (Fig. 1B). Given these articulatory configurations, we hypothesized that infants in experiment 2 would fail to discriminate the Hindi contrast, whereas those in experiment 3 would show successful discrimination. Ultrasound imaging was used to verify that the teethers resulted in the appropriate manipulation of tongue placement in experiments 2 and 3.

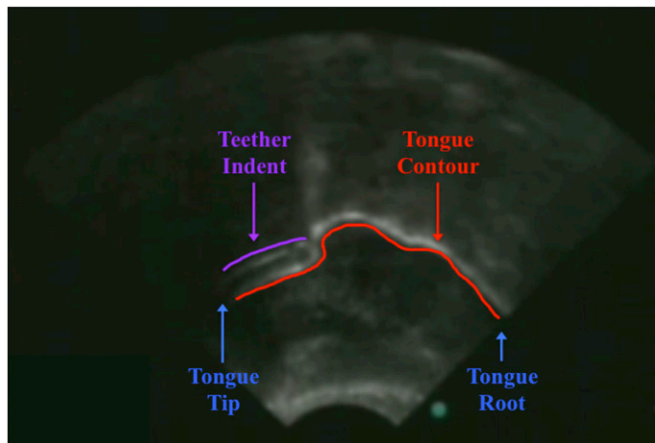
In experiments 1–3, 6-mo-old English-learning infants were tested in a standard alternating/nonalternating (Alt/NAlt) procedure in which looking time to a checkerboard was the dependent variable (39). Infants were presented with two types of trials: those in which tokens from the dental and retroflex phonetic categories alternated in presentation (Alt trials), and those in which tokens from the same speech sound category were repeated for the duration of the trial (NAlt trials). In this design, significantly longer looking time to the Alt over the NAlt trials is taken as evidence that infants discriminate between these two sound categories.

### Experiment 1

Twenty-four English-learning 6-mo-old infants participated in experiment 1. Looking time data were analyzed across the four trials of each type (four Alt, four NAlt). Following Yeung and Werker (40), looking time data were analyzed in pairs of trials, where each pair contained one Alt and one NAlt trial: Pair 1 included the first two trials of the study, pair 2 included the third and fourth trials, and so on; this allowed us to account for any changes in looking time across the series of trials (see *Methods* for methodological details). A 2 (Trial Type)  $\times$  4 (Pair) repeated-measures ANOVA was performed on the looking times, using the within-subjects factors of Trial Type (Alt or NAlt) and Pair (first, second, third, or fourth). The main effect of Pair was significant,  $F(3, 69) = 9.84, P < 0.001, \eta_p^2 = 0.30$ , indicating that infants' looking time significantly declined across the four pairs of trials, as is standard in familiarization or habituation looking time paradigms [the linear contrast was significant,  $F(1, 23) = 23.91, P < 0.001, \eta_p^2 = 0.51$ ]. Further, there was a significant main effect of Trial Type,  $F(1, 23) = 4.32, P = 0.049, \eta_p^2 = 0.16$ , and no interaction with Pair,  $F(3, 69) = 1.63, P = 0.19, \eta_p^2 = 0.066$ , suggesting that the difference in looking time between the two types of trials did not significantly differ across the four pairs. Follow-up investigation of looking time means for the significant Trial Type effect showed that infants looked longer during Alt trials ( $M = 9,369.17$  ms,  $SD = 4,033.06$ ) than NAlt trials ( $M = 8,542.29$  ms,  $SD = 4,053.58$ ). The results from experiment 1 replicate previous findings showing that 6-mo-old English-learning infants are able to discriminate the nonnative Hindi /d/-/ɖ/ contrast. This justified the use of the same Alt–NAlt procedure in the following two looking time experiments.

### Teether Validation

To validate the choice of teething toys used in the sensorimotor manipulations in experiments 2 and 3, we ran an ultrasound imaging study in which infants' tongue placements were recorded while infants used two different teethers. One teether was hypothesized to impact the tip and blade of the tongue when inserted into the infant's mouth, given its flat, planar shape (and will henceforth be called the "flat" teether; Fig. 1A). A second



**Fig. 2.** Ultrasound image from live recording (flat teether in the infant's mouth). The infant's tongue contour ("profile" of the infant's tongue) is the thick, curved white line in the middle of the sonogram (underlined in red); the tongue tip is to the left, and the tongue root is to the right (denoted in blue). At the tongue tip, notice the indent in the tongue contour due to placement of the flat teether (shown in purple).

teether comprised a soft U-shaped silicone pad that fit between the infant's gums but due to the U-shape was not expected to affect the placement of the tongue tip (and will henceforth be called the "gummy" teether; Fig. 1*B*). A sample of three infants' tongues was imaged using a portable ultrasound machine in three situations: (i) while they had no teether in the mouth, (ii) while they had the flat teether in the mouth, and (iii) while they had the gummy teether in the mouth (see Fig. 2 for an example ultrasound image).

Infant ultrasound data showing smoothing spline (SS) ANOVAs (41) are separated in pairs of three teether manipulations (Fig. 3; see Figs. S1–S3 for complete ultrasound data). The tongue contours depicted in Fig. 3 show the shape of the infant's tongue profile with the different teethers in place, or with no teether; the left end of each tongue contour represents the tongue tip, and the right end of the contour represents the tongue root. As described in *Methods*, we used 95% Bayesian confidence intervals (CIs) to determine whether tongue tip placement was significantly different in the three teether manipulations. There was no significant difference in tongue tip placement between the no teether (gray lines) and gummy teether (teal lines) manipulations in all three infants; however, there were significant differences in tongue tip/blade contours between the no teether and flat teether (pink lines) manipulations as well as between the gummy teether and flat teether manipulations. These ultrasound imaging results confirm that the flat teether, but not the gummy teether, selectively impacts the position of the infants' tongues.

### Experiment 2

In experiment 2, 24 English-learning 6-mo-old infants were tested in an identical Alt–NAlt test procedure to that used in

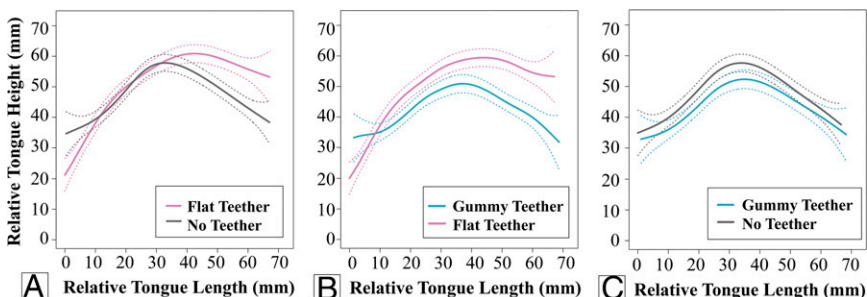
experiment 1. The only modification was that a caregiver held the flat teething toy in the infant's mouth for the entire duration of the study, thereby impeding the movement and controlling the placement of the infant's tongue tip. With this manipulation, experiment 2 tested the hypothesis that temporarily disabling the tongue's movement would impede auditory discrimination of speech sounds that require a distinction in tongue placement during production.

Looking time data were analyzed across the four trials of each type (four Alt, four NAlt) in the four pairs of trials as described earlier. A 2 (Trial Type)  $\times$  4 (Pair) repeated-measures ANOVA was performed on the looking times, using the within-subjects factors of Trial Type (Alt or NAlt) and Pair (1, 2, 3, or 4). There was a main effect of Pair,  $F(3, 69) = 8.61$ ,  $P < 0.001$ ,  $\eta_p^2 = 0.27$ , indicating that infants' looking time significantly declined across the four pairs of trials [the linear contrast was significant,  $F(1, 23) = 17.53$ ,  $P < 0.001$ ,  $\eta_p^2 = 0.43$ ]. However, there was no main effect of Trial Type,  $F(1, 23) = 0.011$ ,  $P = 0.92$ ,  $\eta_p^2 < 0.001$ , and no interaction with Pair,  $F(3, 69) = 1.17$ ,  $P = 0.33$ ,  $\eta_p^2 = 0.048$ , suggesting that the pattern of looking time to the two types of trials was similar across the four pairs of trials. Across the four trial pairs, infants did not look any longer during Alt trials ( $M = 9,846.67$  ms,  $SD = 4,099.99$ ) compared with NAlt trials ( $M = 9,803.54$  ms,  $SD = 4,311.07$ ) while the flat teether was in the mouth (Fig. 4). The results from experiment 2 provide evidence that sensorimotor information influences speech perception of nonnative speech sounds in 6-mo-old infants: **Infants who were given a flat teether that temporarily prevented movement of their tongue tips failed to show evidence of discriminating a phonetic contrast that differed in tongue tip placement.**

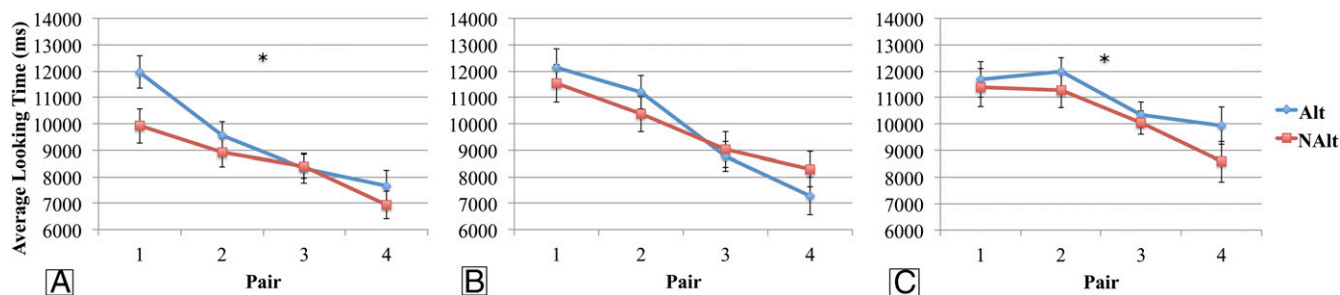
### Experiment 3

Experiment 3 was designed to test the specificity of the teether effect found in experiment 2 and to rule out the possibility that the flat teether drew attention away from the task. Experiment 3 was identical to experiment 2, except that infants were given the gummy teether, which did not impede the movement or placement of the tongue tip. A caregiver held the gummy teether in the infant's mouth for the entire duration of the study. It was hypothesized that discrimination of the phonetic contrast would remain intact with this manipulation. Twenty-four English-learning 6-mo-old infants were tested.

Looking time data were analyzed across the four trials of each type (four Alt, four NAlt) in the four pairs of trials as described above. A 2 (Trial Type)  $\times$  4 (Pair) repeated-measures ANOVA was performed on the looking times, using the within-subjects factors of Trial Type (Alt or NAlt) and Pair (1, 2, 3, or 4). There was a main effect of Pair,  $F(1.96, 45.05) = 3.77$ ,  $P = 0.031$ ,  $\eta_p^2 = 0.14$  [the linear contrast was significant,  $F(1, 23) = 5.14$ ,  $P = 0.033$ ,  $\eta_p^2 = 0.18$ ], meaning infants' looking time significantly declined across the four pairs of trials. (We used Greenhouse Geisser correction on df for violation of sphericity assumption.) Further, there was a significant main effect of Trial Type,  $F(1, 23) = 5.26$ ,  $P = 0.031$ ,  $\eta_p^2 = 0.19$ , and no interaction with Pair,  $F(3, 69) = 0.37$ ,  $P = 0.77$ ,  $\eta_p^2 = 0.016$ , suggesting that the difference in looking time between the two types of trials did not significantly differ



**Fig. 3.** Ultrasound results. One infant's tongue contours comparing (A) flat teether (in pink) versus no teether (in gray), (B) gummy teether (in teal) versus flat teether (in pink), and (C) gummy teether (in teal) versus no teether (in gray). As in Fig. 2, tongue tip is to the left, and the tongue root is to the right. Dotted lines denote 95% CIs.



**Fig. 4.** Looking time averages during test trials in experiments 1–3. (A) Experiment 1 average looking time to each type of trial (Alt in blue and NAlt in red) for the four pairs of trials. (B) Experiment 2 average looking time to each type of trial (Alt in blue and NAlt in red) for the four pairs of trials. (C) Experiment 3 average looking time to each type of trial (Alt in blue and NAlt in red) for the four pairs of trials. Error bars denote SEMs with between-subjects variability removed. An asterisk indicates Trial Type main effect significance at  $P < 0.05$ .

across the four pairs. Across the four pairs of trials, infants looked longer during Alt trials ( $M = 10,986.46$  ms,  $SD = 4,778.55$ ) compared with NAlt trials ( $M = 10,324.38$  ms,  $SD = 4,640.78$ ) (Fig. 4). The results from experiment 3 indicate that teething toys do not generally disrupt performance on the nonnative speech sound discrimination task; instead, while using a gummy teether that did not impede tongue movement, 6-mo-old English-learning infants successfully discriminated the Hindi /d/-/q/ contrast, as did infants in experiment 1. These findings provide corroborating evidence that the teether effect seen in experiment 2 is due to the particular shape of teether used, one that selectively impacted the infant's placement (and movement) of the tongue tip and blade. To graphically illustrate the difference in discrimination across experiments 1–3, average looking time difference scores (Alt – NAlt) for each experiment are plotted in Fig. 5.

### General Discussion

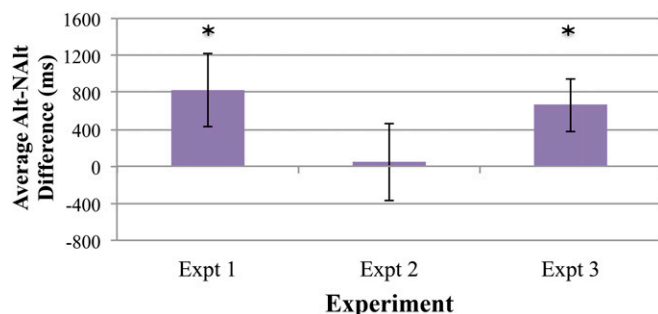
Our work reveals that in early infancy, auditory discrimination of a nonnative speech sound contrast is selectively impaired by temporarily disabling the tongue movement relevant for the production of that contrast. These findings demonstrate a sensorimotor influence on speech perception in preverbal infants. Importantly, the articulatory-induced impairment in discrimination was evident in infants who had no experience listening to these sounds or seeing these sounds being produced, as they comprised a (Hindi) nonnative contrast. This indicates that infants' own articulatory configurations can influence auditory speech perception, even without experiencing the association between the particular visible oral–motor movements and the corresponding sounds.

The study of motor influences on the perception of speech has a long history (42–45), yet the available evidence for a sensorimotor influence on the perception of speech in developmental populations is limited. Our findings not only augment a growing body of work suggesting that infants have the neural architecture to combine auditory and sensorimotor information from early in life (33–35) but also provide the first experimental evidence, to our knowledge, of a sensorimotor influence on auditory perception in infancy. Specifically, these findings show productive influences on perception, even without the accrual of specific experience. We therefore posit that as they acquire the native language, infants use both auditory input to guide speech motor behavior and sensorimotor input to inform speech perception. Our results suggest that both the acoustic percept and sensorimotor information from the articulators are involved in a bidirectional perception–production link in early infancy, thus challenging an “acoustics first” account of such a linkage (i.e., refs. 44, 45). Importantly, the current research cannot distinguish whether there is a single, innate representation for the perception and production of speech as in the strong version of the motor theory of speech perception (43) or whether perception and production are separate but linked, with each guiding and informing the other throughout the period of language acquisition. Nonetheless, the high degree of articulatory–phonetic fidelity

in these data suggests that even if a single, shared representation for the perception and production of speech is not part of the initial architecture, infants are prepared to rapidly establish precise linkages with only minimal perceptual and early productive experience (see refs. 46, 47 for a discussion of the notion of “innately guided learning”). More research is necessary to tease apart these possibilities.

It is informative to compare our results to those reported in a recent neuroimaging study using magnetoencephalography (MEG) by Kuhl et al. (37). Their results showed that 7-mo-old infants' auditory and motor areas respond equally to native and nonnative speech sounds, whereas by 11 mo of age, motor areas show greater activation to nonnative speech, and auditory areas respond more to native speech. Although Kuhl and colleagues interpret their results more within an analysis-by-synthesis view (44, 45) and place a greater emphasis on experience, they also conclude that more research is needed to disentangle the genesis of motor influences on speech perception. Nonetheless, using different approaches and methodologies, our work and that of Kuhl et al. (37) both reveal a crucial role for the motor system in auditory speech perception from very early in development.

Much of the previous research on sensorimotor, articulatory influences on speech perception in infants has concerned the influences on auditory–visual speech perception, rather than on auditory speech perception. As noted in the introduction, infants are sensitive to the match between auditory and visual speech from very early in life (8–11). Moreover, experimentally manipulating the position of infants' articulators to share qualities of the presented sounds modifies 4-mo-olds' performance in an audiovisual speech-matching task: By giving infants either a pacifier, which rounded the lips to share qualities with “oo” sounds, or a teething ring, which stretched the lips to share qualities with “ee” sounds, infants' ability to match heard and seen speech was modified



**Fig. 5.** Alt–NAlt difference score averages. Average difference in looking time between Alt and NAlt trials for each experiment (in ms). Scores greater than zero indicate an overall Alt > NAlt preference. Error bars denote SEM difference, and an asterisk indicates a significant difference (from zero), as reflected in the individual ANOVAs.

(38). From these findings, Yeung and his colleagues have proposed that the information available in auditory and visual speech is mapped onto a common articulatory representation (48). This interpretation is also supported by vocal imitation studies in preverbal infants (8, 49–51), as in all of these studies the infants had access to both the auditory and visual information in speech. However, such findings (*i*) do not address, or consider, the extent to which sensorimotor information influences the perception of auditory speech directly, without visual support, and (*ii*) cannot rule out the possibility that experience with the native language has played a role in establishing specific sensorimotor influences on speech perception. The current studies address both of these points: We not only show that sensorimotor information from the articulators directly shapes auditory speech perception, but by using nonnative rather than native speech contrasts, we show that specific articulatory-to-auditory mappings can occur without previous perceptual (auditory or visual) or productive experiences with the particular speech sounds being studied.

These results have important theoretical implications. They indicate that the establishment of native phonetic categories incorporates the listener's sensorimotor experiences and, therefore, that sensorimotor, articulatory information may be formative in the process of language acquisition. Thus, although sensorimotor information may be modulatory in adult speech perception (32), our results suggest that sensorimotor information is recruited for more than just supplementary purposes in early infancy.

A hypothesis that follows from these findings is that impairment to the articulatory system could be a hindrance to speech perception and language development. The experimental, temporary manipulation of the articulators used in the current experiments interfered with the perception of speech sounds that the infants had never before experienced. This raises the question of whether speech perception is compromised in clinical populations with congenital oral–motor deficiencies or dysmorphologies. At present, these children are given speech therapy to improve their production skills, but no consideration is given to whether their perception is also affected by differences in their productive systems. Although it is the case that patients with either impaired speech production, or those who have never been able to produce spoken language, are capable of perceiving speech sounds (32), the establishment of phonetic categories in early infancy may rely preferentially on access to sensorimotor information. Without access to that information, speech perception development could be compromised and/or achieved via a different route.

The majority of infants experience sensorimotor disruptions fairly regularly, in the use of teething toys, soothers, or thumb-sucking behavior. We are not suggesting that these events would have long-term detrimental effects on language acquisition; it is possible that any unperturbed sensorimotor experience is sufficient to inform speech perception and hence that these common infant behaviors do not negatively impact perceptual development as long as the infant has some experience using their articulators. Alternatively, it is possible that there is a sensorimotor–articulatory perturbation threshold, beyond which perceptual development is compromised or altered. Moreover, infants whose perceptual systems are already deficient—as in the case of infants with mild to moderate hearing loss—may be at greater risk for speech perception difficulties if they also experience regular articulatory perturbations.

## Conclusions

Sensorimotor information from the articulators selectively affects speech perception in 6-mo-old infants even without productive or perceptual experience with the speech sounds. These findings suggest that a link between the articulatory–motor and speech perception systems may be more direct than previously thought and is available even before infants accrue experience producing speech sounds themselves. The present work provides a foundation for the reconsideration of theories concerning the processing of speech during language acquisition: Such theories must account for the influence of sensorimotor information on speech perception and determine the consequences or advantages of such a linkage as infants acquire the native language.

## Methods

**Participants.** The participants in experiment 1 were 24 infants (12 male; mean age, 6 m, 22 d; range, 6 m, 5 d to 7 m, 25 d), with a 97.50% average exposure to English (range, 90–100% English exposure). Data from three additional infants were not included due to fussiness ( $n = 2$ ) and parental interference ( $n = 1$ ). The participants in the ultrasound study were three infants (two males; mean age, 7 m, 10 d; range, 6 m, 23 d to 7 m, 25 d). The participants in experiment 2 were 24 infants (12 male; mean age, 6 m, 30 d; range, 6 m, 7 d to 7 m, 25 d), with a 98.92% average exposure to English (range, 90–100% English exposure). Data from three additional infants were not included due to fussiness ( $n = 1$ ), equipment failure ( $n = 1$ ), and parental interference ( $n = 1$ ). The participants in experiment 3 were 24 infants (12 male; mean age, 6 m, 27 d; range, 6 m, 6 d to 7 m, 25 d), with 97.63% average exposure to English (range, 90–100% English exposure). Data from eight additional infants were not included due to fussiness ( $n = 3$ ), equipment failure ( $n = 2$ ), and parental interference ( $n = 3$ ). Infants' parents were contacted through a database of families; parents and infants were originally recruited for this database at a local maternity hospital shortly after the infants' birth. Parents gave written consent for their infant's participation before the study began. After the study, infants received a t-shirt and were awarded a certificate as a thank you for participation. This research was approved by University of British Columbia Ethics Certificate B95-0023/H95-80023 (University of British Columbia Behavioral Research Ethics Board).

### Experiments 1–3.

**Stimuli.** The auditory stimuli used in experiments 1–3 were recorded by a female native speaker of Hindi. Stimuli included a set of /*ɖa*/ and /*ɖa*/ syllable tokens recorded in infant-directed speech, where each syllable was spoken in a triplet (i.e., /*ɖa*/ /*ɖa*/ /*ɖa*/); the middle token was spliced from each triplet and used in the final audio files. Each syllable token was analyzed for pitch, duration, and amplitude consistency. Three unique tokens of /*ɖa*/ and three unique tokens of /*ɖa*/ were used to create the eight 20-s auditory stimuli streams described below, and each stream contained 10 tokens. On average, the tokens were 900 ms in duration and 70 dB in amplitude, and the stimulus onset interval was 2 s.

**Apparatus.** Infants' looking time data were collected using a Tobii 1750 eye-tracking system, consisting of a PC-run monitor (34 cm × 27.5 cm screen) that both presented the visual stimuli and captured the infants' gaze information (using Tobii Clearview software) and a Macintosh desktop that controlled the stimuli presentation. The eye tracker collected gaze information every 20 ms, and the area of interest included the full screen (1024 × 768 pixels, or 34 × 27.5 cm). Looking time data were analyzed using a custom Excel script.

**Procedure.** The Alt/NAIt sound presentation task in experiments 1–3 is a commonly used procedure to test speech sound discrimination in infants (39, 40, 52, 53). Eight 20-s stimuli streams were created: Four Alt streams contained five presentations each of /*ɖa*/ and /*ɖa*/ tokens in a random order, and four NAIt streams contained 10 presentations of either /*ɖa*/ or /*ɖa*/ tokens. These auditory streams were presented to the infant while a checkerboard was shown on the computer screen. Each infant experienced a fixed series of the eight trials, which were ordered in such a way that every other trial was Alt (and vice versa). The order of the first stimulus presented was counterbalanced across infants.

Infants were seated on their parent's lap facing the eye-tracking monitor in a dimly lit, sound-attenuated room. Parents were asked not to speak to their infants and listened to masking music over headphones. The eye tracker was calibrated with a five-point display. Before the onset of each trial, infants' attention was drawn to the center of the screen with a colorful looming object. During each trial, the speech sounds were played from two speakers located to the sides of the monitor at a level of 65 dB; the eye tracker recorded looking time to the checkerboard during each trial.

Experiments 2 and 3 used the Alt/NAIt looking time procedure while caregivers held the respective teethers in their infants' mouths for the duration of the experiment: the flat teether (Learning Curve Baby Fruity Teethers) in experiment 2 and the gummy teether (Nuby Gum-Eez First Teethers) in experiment 3. Infants' level of preoccupation with the two teethers was coded offline after the study (see *Data Analyses* for distractibility results). Infants whose parents dropped the teether at any point during the study were excluded from the final sample (experiment 2,  $n = 1$ ; experiment 3,  $n = 2$ ).

### Ultrasound Recording.

**Apparatus.** The equipment used to record ultrasound images included a portable ultrasound machine, the two teething toys (flat and gummy teethers), and video recording software. A Sonosite Titan portable ultrasound machine was used, together with a 5–8 MHz Sonosite C-11 transducer with a 90° field of view and a depth of 8.2 cm. All recordings were done using the Pen setting. Ultrasound videos were captured on a separate PC computer using a Canopus TwinPact100 converter from the Sonosite ultrasound machine.

**Procedure.** Each infant was seated on his or her parent's lap in a comfortable, upright position. No stimuli were presented during the ultrasound sessions. The experimenter placed the transducer under the infant's chin until proper placement was achieved, when the tongue tip was visible on the Sonosite screen. As soon as the infant was comfortable with the transducer, a second experimenter began the video recording. The experimenter held the transducer in place for 10–20 s, or as long as the infant remained comfortable, to obtain an ultrasound recording of the tongue contour with no teether in the mouth.

The parent placed one of the two teethers into the infant's mouth. As soon as the infant was comfortable with the teether, the experimenter again placed the transducer under the chin and held it there until 10–20 s of ultrasound recording had passed (see Fig. 2 for an image of an infant in ultrasound recording with the flat teether in the mouth). Finally, the second teether was placed in the mouth and ultrasound images were again recorded for a period of 10–20 s. The ultrasound recording sessions lasted different lengths of time, depending on the cooperation of the infant and ease of transducer placement under the chin.

**Data analysis.** Twelve separate frames from each of the three teether manipulations (from each infant) were spliced from the video recordings into jpeg images; images were chosen based on the clarity of the image and tongue contour (see Fig. 2 for an example image). EdgeTrak software (54)

was then used to extract the ( $x$ ,  $y$ ) coordinates of the infants' tongue contours; 30 coordinate points were extracted for each frame image.

SS ANOVAs (41) were used to analyze the ultrasound images collected for the three infants in custom R scripts. SS ANOVAs determine whether the shapes of two curves are significantly different from each other and are thus ideal for analyzing ultrasound images of tongue contours during the different teether manipulations. SS ANOVAs return 95% Bayesian CIs to specify whether and where two curves differ from each other. Because we were most interested in the placement of the tongue tip, the area of interest for the SS ANOVA analyses concerned only the front portion (tip) of the tongue, which was always the left side of the tongue contour. Statistical analyses were performed on tongue contours separately for each infant.

**ACKNOWLEDGMENTS.** We thank Bryan Gick, Noriko Yamane, and Melinda Hejil for their assistance in obtaining ultrasound imaging data; Eric Vatikiotis-Bateson for his help recording the auditory stimuli used in experiments 1–3; Yaachna Tangri, Karyn Kraemer, Savannah Nijeboer, and Aisha Ghani for their support in recording stimuli, recruiting infants, and coding data; and H. Henny Yeung for insightful discussions. This work was supported by Eunice Kennedy Shriver National Institute of Child Health & Human Development, National Institutes of Health Grant R21HD079260 and Natural Sciences and Engineering Research Council of Canada Grant DG 81103 (to J.F.W.).

- Eimas PD, Siqueland ER, Jusczyk P, Vigorito J (1971) Speech perception in infants. *Science* 171(3968):303–306.
- Kuhl PK (2004) Early language acquisition: Cracking the speech code. *Nat Rev Neurosci* 5(11):831–843.
- Werker JF, Gervain J (2013) Speech perception in infancy: A foundation for language acquisition. *The Oxford Handbook of Developmental Psychology*, ed Zelazo PD (Oxford University Press, New York), pp 909–925.
- Werker JF, Tees RC (1984) Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behav Dev* 7(1):49–63.
- Kuhl PK, Williams KA, Lacerda F, Stevens KN, Lindblom B (1992) Linguistic experience alters phonetic perception in infants by 6 months of age. *Science* 255(5044):606–608.
- Kuhl PK, et al. (2006) Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Dev Sci* 9(2):F13–F21.
- Narayan CR, Werker JF, Beddor PS (2010) The interaction between acoustic salience and language experience in developmental speech perception: Evidence from nasal place discrimination. *Dev Sci* 13(3):407–420.
- Kuhl PK, Meltzoff AN (1982) The bimodal perception of speech in infancy. *Science* 218(4577):1138–1141.
- MacKain K, Studdert-Kennedy M, Spieker S, Stern D (1983) Infant intermodal speech perception is a left-hemispheric function. *Science* 219(4590):1347–1349.
- Patterson ML, Werker JF (2003) Two-month-old infants match phonetic information in lips and voice. *Dev Sci* 6(2):191–196.
- Pons F, Lewkowicz DJ, Soto-Faraco S, Sebastián-Gallés N (2009) Narrowing of intersensory speech perception in infancy. *Proc Natl Acad Sci USA* 106(26):10598–10602.
- Kent RD (1992) The biology of phonological development. *Phonological Development: Models, Research, Implications*, eds Ferguson CA, Menn L, Stoel-Gammon C (York Press, Timonium, MD), pp 65–90.
- Vihman MM (1996) *Phonological Development: The Origins of Language in the Child* (Blackwell Publishing, Cambridge, MA).
- Oller DK (1980) The emergence of the sounds of speech in infancy. *Child Phonology, Volume 1: Production*, eds Yeni-Komshian GH, Kavanagh JF, Ferguson CA (Academic, New York), pp 93–112.
- Mampe B, Friederici AD, Christophe A, Wermke K (2009) Newborns' cry melody is shaped by their native language. *Curr Biol* 19(23):1994–1997.
- de Boysson-Bardies B, Vihman MM (1991) Adaptation to language: Evidence from babbling and first words in four languages. *Language: Journal of the Linguistic Society of America* 67(2):297–319.
- Wilson SMW, Saygin AP, Sereno MI, Iacoboni M (2004) Listening to speech activates motor areas involved in speech production. *Nat Neurosci* 7(7):701–702.
- Pulvermüller F, et al. (2006) Motor cortex maps articulatory features of speech sounds. *Proc Natl Acad Sci USA* 103(20):7865–7870.
- Bradlow AR, Pisoni DB, Akahane-Yamada R, Tohkura Y (1997) Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *J Acoust Soc Am* 101(4):2299–2310.
- Houde JF, Jordan MI (1998) Sensorimotor adaptation in speech production. *Science* 279(5354):1213–1216.
- Sams M, Möttönen R, Sihvonen T (2005) Seeing and hearing others and oneself talk. *Brain Res Cogn Brain Res* 23(2–3):429–435.
- Ito T, Tiede M, Ostry DJ (2009) Somatosensory function in speech perception. *Proc Natl Acad Sci USA* 106(4):1245–1248.
- D'Ausilio A, et al. (2009) The motor somatopy of speech perception. *Curr Biol* 19(5):381–385.
- Möttönen R, Watkins KE (2009) Motor representations of articulators contribute to categorical perception of speech sounds. *J Neurosci* 29(31):9819–9825.
- Oller DK, Eilers RE (1988) The role of audition in infant babbling. *Child Dev* 59(2):441–449.
- Whalen DH, Levitt AG, Goldstein LM (2007) VOT in the babbling of French- and English-learning infants. *J Phonetics* 35(3):341–352.
- Kent RD (1984) Psychobiology of speech development: Coemergence of language and a movement system. *Am J Physiol* 246(6 Pt 2):R888–R894.
- DePaolis RA, Vihman MM, Keren-Portnoy T (2011) Do production patterns influence the processing of speech in prelinguistic infants? *Infant Behav Dev* 34(4):590–601.
- Majorano M, Vihman MM, DePaolis RA (2014) The relationship between infants' production experience and their processing of speech. *Lang Learn Dev* 10(2):179–204.
- Saur D, et al. (2008) Ventral and dorsal pathways for language. *Proc Natl Acad Sci USA* 105(46):18035–18040.
- Hickok G, Poeppel D (2007) The cortical organization of speech processing. *Nat Rev Neurosci* 8(5):393–402.
- Hickok G, Houde J, Rong F (2011) Sensorimotor integration in speech processing: Computational basis and neural organization. *Neuron* 69(3):407–422.
- Perani D, et al. (2011) Neural language networks at birth. *Proc Natl Acad Sci USA* 108(38):16056–16061.
- Dehaene-Lambertz G, et al. (2006) Functional organization of perisylvian activation during presentation of sentences in preverbal infants. *Proc Natl Acad Sci USA* 103(38):14240–14245.
- Imada T, et al. (2006) Infant speech perception activates Broca's area: A developmental magnetoencephalography study. *Neuroreport* 17(10):957–962.
- Mahmoudzadeh M, et al. (2013) Syllabic discrimination in premature human infants prior to complete formation of cortical layers. *Proc Natl Acad Sci USA* 110(12):4846–4851.
- Kuhl PK, Ramirez RR, Bosseler A, Lin JF, Imada T (2014) Infants' brain responses to speech suggest analysis by synthesis. *Proc Natl Acad Sci USA* 111(31):11238–11245.
- Yeung HH, Werker JF (2013) Lip movements affect infants' audiovisual speech perception. *Psychol Sci* 24(5):603–612.
- Best C, Jones C (1998) Stimulus-alternation preference procedure to test infant speech discrimination. *Infant Behav Dev* 21(1):295.
- Yeung HH, Werker JF (2009) Learning words' sounds before learning how words sound: 9-month-olds use distinct objects as cues to categorize speech information. *Cognition* 113(2):234–243.
- Davidson L (2006) Comparing tongue shapes from ultrasound imaging using smoothing spline analysis of variance. *J Acoust Soc Am* 120(1):407–415.
- Lieberman AM, Cooper FS, Shankweiler DP, Studdert-Kennedy M (1967) Perception of the speech code. *Psychol Rev* 74(6):431–461.
- Lieberman AM, Mattingly IG (1985) The motor theory of speech perception revised. *Cognition* 21(1):1–36.
- Halle M, Stevens KN (1962) Speech recognition: A model and a program for research. *IRE Transactions on Information Theory* 8(2):155–159.
- Stevens KN, Halle M (1967) Remarks on analysis by synthesis and distinctive features. *Models for the Perception of Speech and Visual Form*, ed Walther-Dunn W (MIT Press, Cambridge, MA), pp 88–102.
- Gould JL, Marler P (1987) Learning by instinct. *Sci Am* 256(1):74–85.
- Jusczyk PW, Bertoncini J (1988) Viewing the development of speech perception as an innately guided learning process. *Lang Speech* 31(Pt 3):217–238.
- Guellai B, Streri A, Yeung HH (2014) The development of sensorimotor influences in the audiovisual speech domain: Some critical questions. *Front Psychol* 5:812.
- Kuhl PK, Meltzoff AN (1996) Infant vocalizations in response to speech: Vocal imitation and developmental change. *J Acoust Soc Am* 100(4 Pt 1):2425–2438.
- Chen X, Striano T, Rakoczy H (2004) Auditory-oral matching behavior in newborns. *Dev Sci* 7(1):42–47.
- Coulon M, Hemimou C, Streri A (2013) Effects of seeing and hearing vowels on neonatal facial imitation. *Infancy* 18(5):782–796.
- Maye J, Werker JF, Gerken L (2002) Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition* 82(3):B101–B111.
- Mattock K, Molnar M, Polka L, Burnham D (2008) The developmental course of lexical tone perception in the first year of life. *Cognition* 106(3):1367–1381.
- Li M, Kambhampettu C, Stone M (2005) Automatic contour tracking in ultrasound images. *Clin Linguist Phon* 19(6–7):545–554.

# Supporting Information

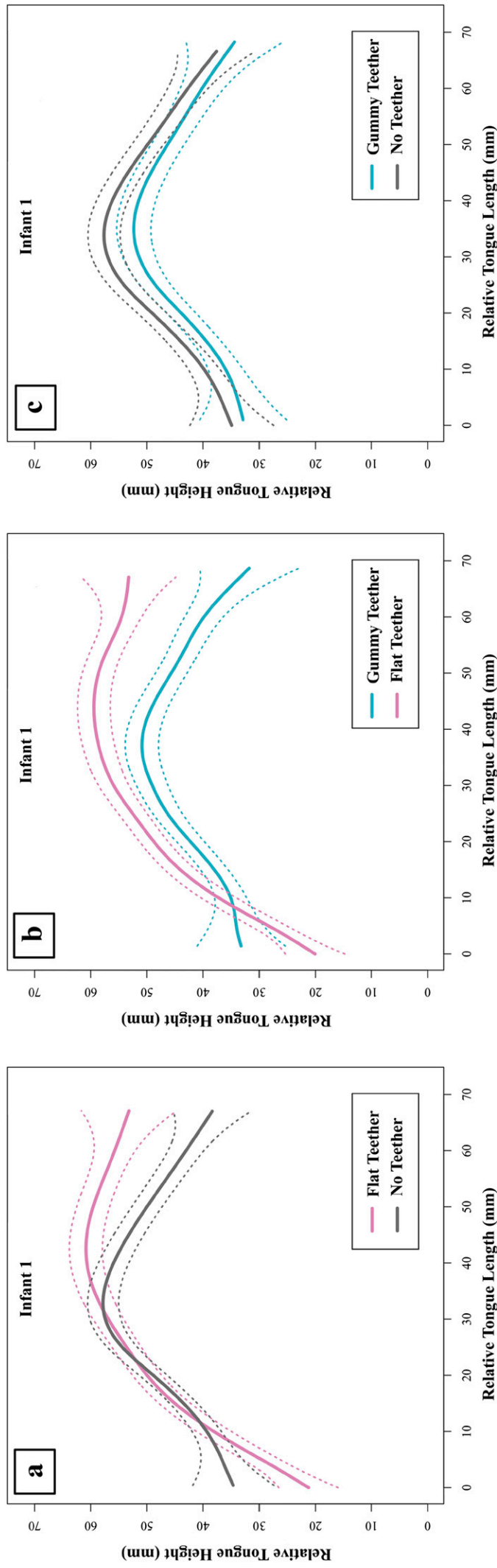
Bruderer et al. 10.1073/pnas.1508631112

## Data Analyses

**Analysis S1, Teether Distractibility Results.** To ensure that attention to the auditory stimuli as measured by looking time in this task was mediated by the position of the tongue and not by the teethers differentially interfering with attention to the speech sounds in experiments 2 and 3, infants' level of preoccupation with the two teethers was coded offline after the study. A coder who was blind to the hypotheses of the work rated each infant on a scale of 1 (very distracted by the teether) to 7 (not at all distracted by the teether). The difference in distractibility was not significantly different between infants in experiment 2 who used the flat

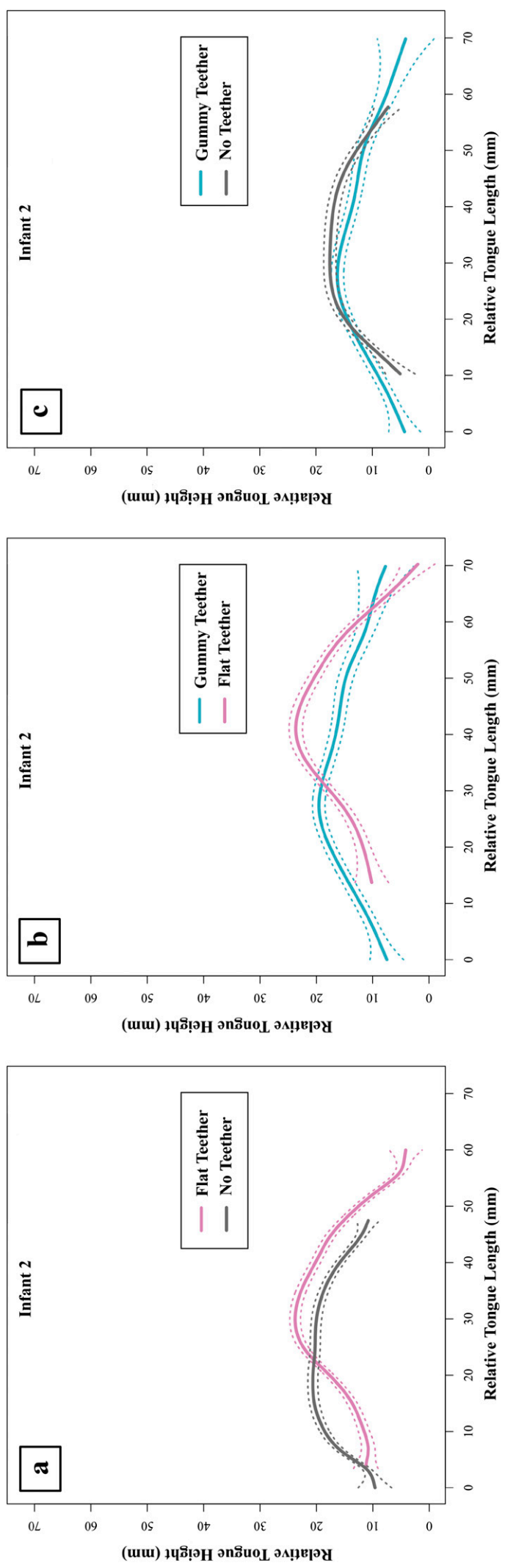
teether (distractibility average score,  $M = 4.92$ ,  $SD = 1.59$ ) and infants using the gummy teether in experiment 3 (distractibility average score,  $M = 5.75$ ,  $SD = 1.48$ ),  $t(46) = -1.88$ ,  $P = 0.066$ , 95% CI of the difference  $[-1.73, 0.06]$ .

To more directly assess differences in attention across experiments 2 and 3, overall looking time in the two experiments was compared. Overall looking time was not significantly different between experiment 2 ( $M = 9,825.10$  ms,  $SD = 4,084.39$ ) and experiment 3 ( $M = 10,655.42$  ms,  $SD = 4,656.78$ ),  $t(46) = -0.66$ ,  $P = 0.52$ , 95% CI of the difference  $[-3,357.38, 1,714.75]$ .

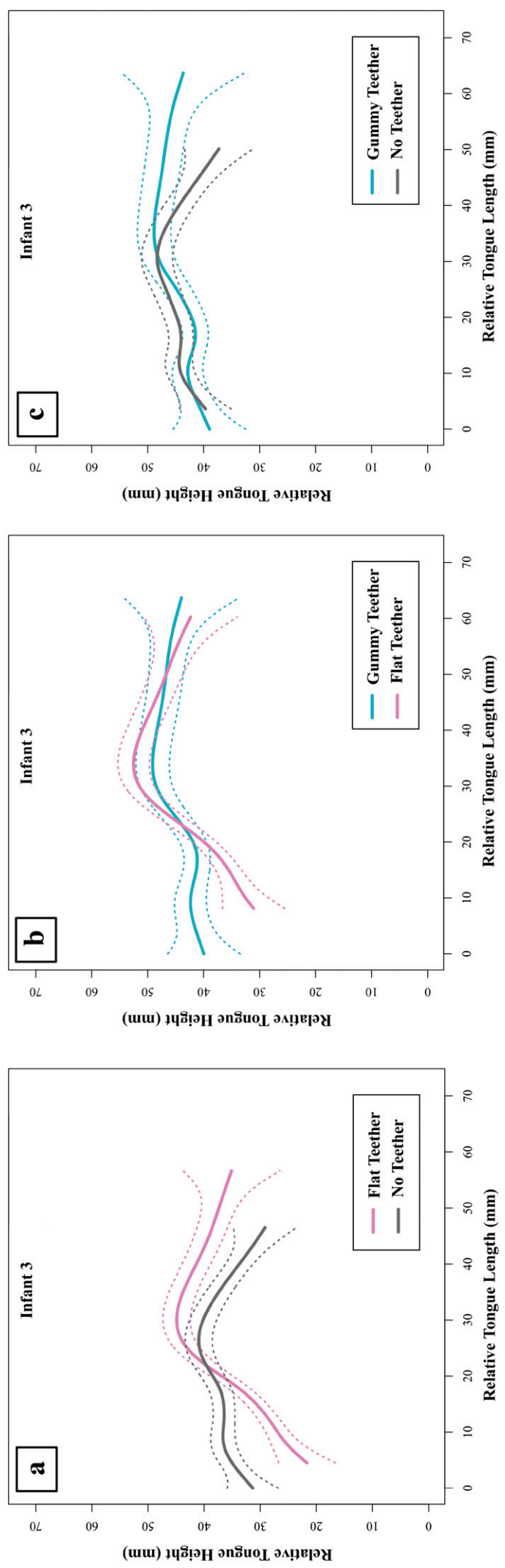


**Fig. S1.** Infant 1 ultrasound results. Infant 1's tongue contours were compared using SS ANOVAs (45). Comparisons include (A) flat teether (in pink) versus no teether (in gray), (B) gummy teether (in teal) versus flat teether (in pink), and (C) gummy teether (in teal) versus no teether (in gray). Tongue tip is to the left, and tongue root is to the right; area of interest is the tongue tip. Dotted lines denote 95% CIs. These data are included in the main text document and are also included here for comparison with infant 2 (Fig. S2) and infant 3 (Fig. S3).





**Fig. S2.** Infant 2 ultrasound results. Infant 2's tongue contours were compared using SS ANOVAs (45). Comparisons include (A) flat teether (in pink) versus no teether (in gray), (B) gummy teether (in teal) versus flat teether (in pink), and (C) gummy teether (in teal) versus no teether (in gray). Tongue tip is to the left, and tongue root is to the right; area of interest is the tongue tip. Dotted lines denote 95% CIs.



**Fig. 53.** Infant 3 ultrasound results. Infant 3's tongue contours were compared using SS ANOVAs (45). Comparisons include (A) flat teether (in pink) versus no teether (in gray), (B) gummy teether (in teal) versus flat teether (in pink), and (C) gummy teether (in teal) versus no teether (in gray). Tongue tip is to the left, and tongue root is to the right; area of interest is the tongue tip. Dotted lines denote 95% CIs.