

Infants' Listening Preferences: Baby Talk or Happy Talk?

Leher Singh and James L. Morgan

*Department of Cognitive and Linguistic Sciences
Brown University*

Catherine T. Best

*Department of Psychology
Wesleyan University*

The most robust finding on infants' listening preferences has been widely characterized as a preference for baby talk (BT) over adult-directed speech (ADS). Although prosodic modifications characteristic of BT also convey positive affect, differences in affect across BT and ADS speech registers have not been controlled in previous studies. This set of experiments sought to elucidate the basis for 6-month-olds' listening preference by independently manipulating affect and speech register. When affect was held constant, no preference for any speech register was observed. Moreover, when ADS stimuli presented more positive affect than BT stimuli, infants' preferences followed the positive affect. Higher and more variable pitch was neither necessary nor sufficient for determining infants' preferences, although pitch characteristics may modulate affect-based preferences. The BT preference is thus attributable to a more general preference for speech that imparts relatively positive affect, a preference perhaps ascribable to a preexisting general-purpose mechanism opportunistically exploited by language.

From birth onward, infants display a range of preferences for auditory stimuli. As early as the first month of life, infants prefer to listen to speech rather than nonspeech (Cooper & Aslin, 1994), mothers' speech rather than other females' speech (DeCasper

Supplementary materials to this article are available on the World Wide Web at <http://www.infancyarchives.com>.

Requests for reprints should be sent to Leher Singh, Department of Cognitive and Linguistic Sciences, Brown University, Box 1978, Providence, RI 02912. E-mail: Leher_Singh@Brown.edu

& Fifer, 1987), and native language speech rather than speech in rhythmically distinct non-native languages (Mehler, Dupoux, Nazzi, & Dahaene-Lambertz, 1996; Moon, Cooper & Fifer, 1993; Nazzi, Bertoncini, & Mehler, 1998). Perhaps the most robust and well-known infant listening preference is what has been characterized as a preference for baby talk (BT) over adult-directed speech (ADS; Cooper & Aslin, 1990; Fernald, 1985). Infants have been shown to listen longer to speech in the BT register in both female and male speech (Pegg, Werker, & McLeod, 1992), both unfiltered and low-pass filtered speech (Fernald, 1985; Fernald & Kuhl, 1987), and both native and non-native languages (McRoberts & Whaley, 1995). Nevertheless, the basis for this set of preferences has never been clearly elucidated. In this article, we present a series of studies demonstrating that infants' listening preference is determined neither by speech register nor by simple acoustic properties characteristic of BT, but rather is chiefly due to the relative positive affect of the speech directed toward infants.

Prelinguistic infants typically encounter many types of speech in their linguistic environments. Among the most common is speech in the register variously termed motherese, parentese, infant-directed speech, or, simply, BT. The assortment of labels for this style of speech underscores the difficulty in establishing a term that appropriately characterizes the register. For purposes of transparency, we adopt BT to refer to speech that is prosodically modified by adult speakers when addressing infants. BT incorporates systematic and identifiable modifications not only to prosody, but also to syntax, semantics, and pragmatics (Brown, 1977; Gleitman, Newport, & Gleitman, 1984). Newborn infants have been reported to show a preference for BT over ADS (Cooper & Aslin, 1990), suggesting that substantial post-natal linguistic experience may not be necessary to elicit this preference. These demonstrations of BT preference have fostered speculation that BT may be tailored to suit infants' language processing and learning capacities (Furrow, Nelson, & Benedict, 1979; Snow, 1999).

Among the prosodic modifications characteristic of BT are elevated pitch, increased pitch variation, expanded intonation contours, reduced speech rate, elongated vowels, long pauses between utterances, and greater prosodic repetition (Cruttenden, 1994; Fernald & Kuhl, 1987; Fernald & Simon, 1984; Papoušek, Bornstein, Nuzzo, Papoušek, & Symmes, 1990; Pegg et al., 1992; Trainor, Austin, & Desjardins, 2000). Pragmatic, syntactic, and semantic characteristics of BT include imitation of infants' utterances, repetition, low mean length of utterance (MLU), and reduced syntactic and semantic complexity (Garnica, 1977; Snow, 1972). Cross-linguistic studies have shown that BT is commonly adopted by adults of different linguistic communities (Fernald et al., 1989), although there are cross-cultural differences in the degree of prosodic modification. English-speaking American parents, in particular, show more exaggerated intonation contours than do adults of other cultures (Fernald et al., 1989). Certain features of BT, such as pitch modification, may be absent in a minority of cultures (Bernstein-Ratner & Pye, 1984; Fee & Shaw, 1998).

Observations of typical BT prosody have given rise to hypotheses that infants' listening preferences may be attributable to specific acoustic characteristics, such as greater pitch height, greater pitch variation, or greater incidence of expanded intonation contours. For example, Fernald and Kuhl (1987) attempted to identify specific acoustic parameters of BT that appeal to infants. They isolated amplitude, duration, and fundamental frequency (F_0) features of BT and showed that variation in F_0 contour, combined with rhythmic factors, may be sufficient to elicit 4-month-olds' preferences. However, recent work by Kitamura and Burnham (1998) investigating infants' preferences for BT speech samples showed that preferences failed to appear when pitch characteristics varied but speech affect was held constant across samples. Conversely, infants in their studies did display preferences for different BT samples when pitch characteristics (high vs. low) were held constant but speech affect was allowed to vary. Together, these results suggest that infants' preferences may be predominantly due to functional, affective properties of BT that are correlated with, but not identical to, simple acoustic and prosodic features of BT.

Several theorists have emphasized the adaptive significance of BT (Fernald, 1992; Fisher & Tokura, 1996; McLeod, 1993; Papoušek et al., 1990; Sachs, 1977). In her summary of this literature, Fernald (1989) enumerated three functional roles of BT. First, BT may serve to modulate infants' attention and state of arousal. Acoustically salient properties, such as F_0 modulation (Fernald & Kuhl, 1987), may launch dyadic interaction between child and caregiver.

A second function of BT relates to its potential pedagogical properties. When infants are confronted with the task of identifying relevant units in the speech stream, they may exploit prosodic regularities that serve as markers for syntactic constituents (e.g., Kemler-Nelson, Hirsh-Pasek, Jusczyk, & Wright-Cassidy, 1989). For example, mothers lengthen vowel durations to mark phrase boundaries (Morgan, 1986), and cross-linguistic evidence for acoustic cues to syntactic structure in BT (Fisher & Tokura, 1996) implicates language-general markers for large-scale syntactic units. BT may also assist identification of novel words, as Golinkoff and Alioto (1995) showed. In their study, adults performed better in learning a novel foreign target word when it was embedded in non-native BT than when it was embedded in non-native ADS. Thus, BT contains multiple cues (e.g., phonological, prosodic, phonotactic) from which infants may potentially extract information about basic linguistic structure.

A third function of BT concerns communication of affect. BT provides an exaggerated indication of speaker affect, making emotion easier to identify in BT than in ADS (Fernald, 1989, 1992). The melodic contours of BT effectively communicate emotion independent of linguistic content. Infants are able to discriminate positive and negative emotions from birth (Mastropieri & Turkewitz, 1999) and respond differentially to positive and negative emotion in speech (Fernald, 1993; Papoušek et al., 1990). For example, Fernald (1993) reported that 5-month-old infants display consistent facial expressions in response to emotional BT,

suggesting that a positive-negative affective distinction is intrinsically meaningful to infants. As described earlier, Kitamura and Burnham (1998) found that infants show a bias toward happy, expressive BT over flat BT, even when pitch was equated. When vocal affect was equated, they found no preference for high-pitch BT over low-pitch BT. These findings suggest that infants are sensitive to affective information inherent in BT, independent of pitch, and that they show a perceptual bias toward positive exemplars over neutral exemplars of BT.

In ADS as well, intonation is an effective communicator of speaker affect (Scherer, 1986). Prosodic features that convey positive affect in adult speech are also common to BT (Fernald, 1989, 1992). Given that BT and happy speech share similar prosodic features, it seems likely that BT typically imparts positive affect, an observation noted in numerous studies (Fernald, 1989, 1992, 1993; Kitamura & Burnham, 1998; Papoušek et al., 1990). However, of all the reports of BT-over-ADS preference, none have reported controlling for affective differences between BT and ADS stimuli. This raises the possibility that studies testing infants' preferences for BT over ADS may have confounded affect and register by contrasting positively emotive BT stimuli with more matter-of-fact ADS stimuli. Thus, the BT preference may actually reflect a general preference for positive affect in speech. This in turn might implicate a deep-seated bias toward displays of positive affect by conspecifics (Darwin, 1872).

Vocal affect expression has been documented as a precedent to linguistic expression both in the evolution of the species and the development of the individual (Bloom, 1990; Darwin, 1872). Infants are capable of expressing themselves using affective vocalizations from birth, and they exploit vocal cues to communicate information about their internal state prior to the onset of productive language (Bloom, 1990; Lewis, 1936). Similarly, many other mammalian communication systems incorporate perceptually salient vocal affect cues to signify motivational or intentional processes. Positive affect attracts the attention of conspecifics and elicits an approach reaction from others, as it typically emanates from a nonthreatening source or caregiver (Darwin, 1872). In humans, a bias toward visual expressions of positive affect has been repeatedly demonstrated: Preverbal infants prefer happy faces over neutral or negative faces, providing support for an innate sensitivity to expressions communicated by the human face (Serrano, Iglesias, & Loeches, 1992, 1995). Given the biological significance of the human voice, it is also important to examine positive affective biases in vocalizations.

Although early preference for positive vocalizations has been documented in infants (Kitamura & Burnham, 1998; Papoušek et al., 1990), it is not known whether such a bias might account for the BT preference. If so, infants' preference for positive affect should be expected to override preference for register. If BT preference is attributable to an affective bias, this would suggest that infants' attention is directed to a broader class of sounds during their prelinguistic months, of which BT may constitute a subset. Furthermore, this would indicate that BT

capitalizes on positive affect as a means to recruit infants' attention to BT, rather than the converse.

No studies have systematically investigated the possibility that positive affect may govern infants' perceptual preferences for BT over ADS. However, Trainor et al. (2000) explored the role of affect in adults' productions of BT and ADS. In an effort to characterize determinants of BT prosody, Trainor et al. measured pitch characteristics, tempo, and rhythmic qualities of BT and ADS and found that, when affect was controlled, few acoustic differences remained between BT and ADS. Their study suggests that adults' canonical productions of BT derive from a tendency to communicate affectively in the presence of infants. This conclusion is supported by cross-cultural investigations of BT, which have shown that prosodic variation in BT is increased in cultures that welcome open expression of emotion and reduced in societies where emotional expression is tempered (Trainor et al., 2000).

This set of studies investigated a similar relationship between affect and register in infants' perceptual preference for speech. We expected that infants would show an overall preference for BT over ADS, replicating findings of earlier studies. Beyond this, we were uncertain whether infants would show a preference for positive emotion independent of register, or a preference for BT register independent of emotion. Our stimuli were explicitly designed to manipulate speech affect and register independently while controlling for other aspects of BT that may engage infants' attention. Thus, our studies were designed to determine whether infants' demonstrated listening preference is for baby talk or happy talk.

STIMULUS DESIGN, RECORDING, AND ANALYSES

Our goal was to create a set of stimulus materials in which speech affect (happy, neutral, or sad) and speech register (BT or ADS) varied orthogonally. In this fashion, we could test infants' preferences for register when affect was controlled (Experiment 2), preferences for affect when register was controlled (Experiment 3), or preferences for speech samples when affect and register were coupled in ways other than those typical in previous studies of infants' preferences (Experiments 4 and 5). Although we did not explicitly manipulate pitch height, pitch variation, or incidence of expanded intonation contours in our stimulus design, the set of comparisons we conducted allowed us to assess the contributions of these simple acoustic properties to infant listening preferences as well.

A potential caveat in controlling for affect involves the difficulty in simulating affective styles. Despite efforts to decompose affectively charged speech into its underlying components (Kappas, Hess, & Scherer, 1991; Scherer, 1986), there appears to be little consensus about the acoustic features that distinguish particular emotions. Affect is therefore better characterized by constellations of

parameters than by single, identifiable acoustic features that uniquely specify particular emotions.

Useful parameters for perception of affect include F_0 , intonation, loudness, speech rate, and first and second formant frequencies (Scherer, 1986). Certain emotions, including depression, boredom, guilt, and indifference tend to be conveyed by low fundamental frequencies, narrow fundamental frequency range, low amplitude, and decreased speech rate. More extreme emotions, including happiness, anger, elation, and fear, tend to be conveyed by high F_0 , wide-pitch excursions, high amplitude, and increased speech rate (Scherer, 1986). From these analyses, it appears that simple prosodic parameters may be useful in distinguishing affective intensity rather than affective valence. However, specific intonation contours tend to be correlated with positive versus negative emotions (Fernald, 1993). Moreover, it has been shown that facial expressions can affect the frequency of the second formant in happiness, where the second formant is raised for certain vowels while the speaker is smiling (Tartter & Braun, 1994). Because other facial expressions associated with particular emotions (Ekman & Friesen, 1971) also affect the shape of the vocal tract, it is possible that spectral composition and speech timbre, in combination with intonation contour, are better indicators of emotional valence.

Generating emotional speech that sounds authentic is difficult; as a result, synthesized speech often fails to sound affectively real. Nevertheless, emotional speech simulated by human speakers can be labeled by listeners with a remarkable degree of accuracy and consensus (Banse & Scherer, 1996; Scherer, 1986). Given this, in producing and evaluating stimuli for our studies, it seemed most methodologically sound to rely on convergent evidence from perceptual categorization and acoustic analyses of emotional stimuli rather than to attempt to manipulate acoustic variables to simulate particular affects.

There are inevitable constraints attached to both spontaneous and scripted recordings. Whereas spontaneous recordings capture infants' speech environments more accurately, they introduce extraneous linguistic factors inherent in BT (e.g., reduced sentence complexity, simplified vocabulary, and increased repetition). Given our central concern with teasing apart speech register from speech affect, it was appropriate that only prosody and spectral composition should vary across stimuli, whereas syntactic and semantic information was held constant. Moreover, scripted recordings using an imaginary scenario have been successfully employed in similar studies (Fernald, 1993; Papoušek et al., 1990) that examined infants' listening preferences for positive emotion in speech. Fernald (1993) used both spontaneous and scripted recordings, and no systematic differences in infants' listening preferences were evident in her data.

Stimuli were recorded by a female native English speaker, who was provided with a total of 18 passages, consisting of six sets of target sentences embedded in three emotional contexts each. Each target sentence was embedded in three

different passages, which described happy, neutral, or sad situations designed to induce corresponding emotional states. The target sentence appeared penultimately in each passage (see Appendix for sample passages). The speaker was asked to rehearse each passage several times. After committing each sentence to memory, she was then asked to repeat the content of the positive passages in ADS. She was asked to smile while speaking and to raise the pitch of her voice toward the end of each sentence to simulate positive affect (Tartter & Braun, 1994). After a satisfactory repetition of each sentence, the speaker moved on to the neutral category. For this category, she was asked to talk to someone else as if she were having an ordinary conversation. Finally, the speaker read the sad category. For this category, she was asked to convey sadness vocally and to display a sad facial expression. She was asked to lower her pitch toward the end of the sentence and to express the disappointment inherent in the passage.

After the adult-directed stimuli were generated, the speaker was asked to read the same passages in BT while looking at a photograph of her infant. For happy BT, the speaker was asked to speak to the infant in BT before reading the passage to the infant, while smiling at the photograph as much as possible. She was then asked to speak to the infant in a neutral style, without imparting any specific emotion. For this affect category, she was instructed to speak very slowly and to enunciate each word more carefully to differentiate neutral BT from neutral ADS. Finally, she was asked to read the sad passages to the infant and sound distressed to convey to the infant the level of unhappiness in the passage. She was asked to lower her pitch toward the end of the sentence and to display a sad facial expression when speaking to the infant. The speech samples judged by the experimenter to effectively convey the desired emotions and registers were sampled into sound files that were later concatenated in random order to be rated by adults.

Ten English-speaking adults were presented with each sentence and asked to rate its affect and register. Each adult listener heard each sentence token three times consecutively. There was a 2-sec pause between each of the three repetitions of the same sentence token and a 4-sec pause between different sentence tokens. Using a 7-point Likert scale ranging from 1 (*happy*) to 7 (*sad*), listeners rated the sentence tokens for their emotional intensity and valence. Neutral emotion was represented by the intermediate point on the scale (4). After rating the stimuli for emotional valence, listeners heard the same stimuli played again in a different order and rated each sentence token for register, again according to a 7-point Likert scale, where 1 represented *infant-directed speech* and 7 represented *ADS*.

Responses for emotion were considered correct if the answer matched the intended emotion or if it departed from the target by 1 point on the 7-point scale. For emotion, the average accuracy across all sentences was 89.16%, with correct responses judged as accurate. Responses for register were judged accurate if they matched the target rating or departed from it by 1 point on the 7-point scale. The mean accuracy rating across sentences and subjects for register was 88.75%.

These accuracy ratings pertain to the entire set of six sentences. However, only four sentences were required for the infant study, so the two sentences with the lowest mean accuracy of ratings across emotion and register were eliminated, yielding four target sentences in three emotional styles by two registers each. The four sentences were matched in lexical content across affect and register (i.e., each sentence was spoken in two registers and three affective tones). For these four sentences, mean accuracy of emotion judgment was 95.2%, and mean accuracy of register judgment was 93.13%. To ensure that we had not biased the first group's ratings by providing them with a happy-sad continuum, we asked an additional group of 19 adults to volunteer their own impressions of the emotions captured by the stimuli while rating the stimuli on a 7-point positive-to-negative affect scale. Again, using the criteria described earlier, ratings were highly correlated with the target emotional valence ($r = .91$). The most common descriptors cited for each set of stimuli were "happy" and "excited" for happy sentences, "no emotion" and "matter-of-fact" for neutral sentences, and "sad" or "upset" for sad sentences. Therefore, the affective labels that we applied to the stimuli appear to be suitably accurate.

To derive a measure of affective intensity of emotional stimuli, we asked 10 adults to rate the selected stimuli for intensity. For these ratings, a 7-point intensity was used ranging from 1 (*emotionally flat*) to 7 (*either extremely intense happiness for happy stimuli or extremely intense sadness for sad stimuli*). Happy BT stimuli ($M = 6.9$) were rated as more intense than happy ADS stimuli ($M = 5.7$), $t(9) = 6.98, p < .0001$, and sad BT stimuli ($M = 6.9$) were rated as more intense than sad ADS stimuli ($M = 5.8$), $t(9) = 14.04, p < .0001$. These findings are consistent with Fernald's (1989) analyses of adults' judgments of communicative intent in BT and ADS, in which she found that low-pass-filtered BT communicated the speakers' intent more effectively than ADS.

Acoustic analyses were conducted on the 24 stimulus sentences. For each sentence, calculations were made for duration, mean F_0 , F_0 range, and formant frequencies for various vowels. Durations of each sentence were measured across register and affect; results are shown in Table 1. Durations of BT sentences were significantly longer than those of ADS sentences, $F(1, 22) = 31.4, p < .001$. There was no main effect of affect, $F(2, 21) = 2.65, p = .09$, and no register by emotion interaction, $F(2, 18) = .63, ns$.

Measures of mean F_0 and mean F_0 range for the selected stimuli are shown in Table 2. For mean F_0 , there was a significant main effect of affect, $F(2, 21) = 6.775, p < .01$, a marginal main effect of register $F(1, 22) = 3.77, p < .05$, and a significant Affect \times Register interaction, $F(2, 18) = 17.09, p < .001$. In ADS, happy stimuli had a higher mean F_0 than sad or neutral stimuli, $t(3) = 3.55, p < .05$ and $t(3) = 5.87, p < .01$, respectively. Sad and neutral stimuli did not differ in mean F_0 . In BT, happy stimuli had a higher mean F_0 than neutral and sad stimuli, $t(3) = 6.91, p < .01$ and $t(3) = 55.61, p < .05$, respectively. Sad BT had a

TABLE 1
Duration Measurements

	ADS	BT	M
Happy	1931.40	3151.88	2541.64
Neutral	1595.23	2643.28	2119.26
Sad	2265.57	3782.65	3024.11
M	1930.73	3192.60	

TABLE 2
 F_0 Measurements Across Affect and Register

	M			M Range (Semitones)		
	ADS	BT	Overall	ADS	BT	Overall
Happy	242.81	246.99	244.90	18.99	24.52	21.75
Neutral	188.79	183.77	186.28	13.56	14.18	13.87
Sad	187.94	282.89	235.42	13.33	22.27	17.80
M	206.51	237.88		15.29	20.32	

higher F_0 than neutral BT, $t(3) = 9.38, p < .01$. Across the register, sad BT had a higher F_0 than sad ADS, $t(3) = 31.18, p < .0001$, but there were no mean F_0 differences across the register for happy or neutral stimuli.

To calculate mean F_0 range, raw frequencies were converted to semitones within each utterance.¹ This provides a nonlinear transformation of raw frequency values that corrects for nonlinearity in perception of pitch range across different pitch levels. Mean F_0 range was significantly wider for BT than for ADS, $F(1, 22) = 2.18, p < .05$. There was a main effect of affect for mean F_0 range, $F(2, 21) = 4.31, p < .05$. Across both registers, mean F_0 range for happy speech was significantly wider than for neutral speech, $t(14) = 3.7, p < .01$. There were no significant differences in mean F_0 range between happy and sad, or sad and neutral speech. There were no interactions between affect and register.

First and second formant frequencies were analyzed, but there were no significant differences across register or affect. In general, differences in these acoustic measures are consistent with those reported in the literature to distinguish spontaneous BT from ADS (Fernald, 1993) and those used to distinguish happiness (elation) and sadness (Scherer, 1986). Based on adult ratings and acoustic

¹The number of semitones in the interval between F_{min} and F_{max} is $c \log_2(F_{max}/F_{min})$. For the standard Western diatonic scale, $c = 12$. Hence, two frequencies perceived to be an octave apart (i.e., $F_{max} = 2 F_{min}$) differ by 12 semitones.

measurements, these stimuli were judged to convey the desired register and affective styles effectively.²

EXPERIMENT 1

The objective of this set of studies was to determine whether the BT preference is contingent on the affective qualities conveyed by BT. We explored this hypothesis by measuring infants' preferences for different speech styles using the head-turn preference procedure (HPP). As noted earlier, several previous studies have shown infant preferences for BT over ADS; that is, happy BT over affectively neutral ADS, using similar preference procedures. As an initial step, to validate both our choice of stimuli and our implementation of the HPP, we set out to replicate this finding.

Method

Participants. The participants in this preliminary study were 13 full-term infants (9 boys and 4 girls), recruited from Rhode Island birth records. Infants were 6 months old ($M = 185$ days, range = 170–225). Participants were included regardless of their exposure to languages other than English in their home environments. One individual was eliminated due to inattention and failure to complete the test.

Stimuli and apparatus. A subset of the stimuli described previously was employed in this study, comprising happy BT and neutral ADS target sentences. The happy BT stimuli were higher pitched, more variable in pitch, and slower than the neutral ADS stimuli (see Tables 1 and 2). These stimulus sets resemble the typical stimuli employed in reports of BT preference. Testing was conducted in a sound-treated laboratory room. A three-walled testing booth was set up within the testing room. Each wall of the testing booth was 120 cm wide. A chair was placed at the open end of the testing booth for the parent to sit on with the infant on his or her lap, approximately 110 cm from the front of the booth. Two speakers rested on the floor of either side of the test room, between the testing booth and the walls of the testing room. At the infant's eye level (86 cm above the floor), a yellow light was mounted on the front wall. In addition, a green light was mounted on each of the side walls at the same height. A Panasonic CCTV (model VW-1410) video camera was positioned behind the test booth 12.3 cm above the yellow center light, although only the camera lens was visible from the testing booth. In a control room down a hall from the test room, a Panasonic Time Lapse (AG 6040) video recorder and Panasonic monitor (WV-5410) were connected to the video camera in the test room to display

the infants' responses on the monitor. The intensity of the speech stimuli was set at conversation level (75 dB) using a Realistic sound-level meter.

The apparatus used here conforms in most respects to written accounts of the standard HPP (e.g., Kemler Nelson et al., 1995). However, one important change was introduced: Whereas most versions of the HPP have a coder in the experimental room who observes the infant through a peephole, we used a video camera so that the infant could be observed from a remote location. The presence of the coder in the testing room where he or she may be able to hear some of the stimuli has given rise to concerns of experimenter bias (Pinto, Fernald, McRoberts, & Cole, 1999). By physically removing the coder from the testing room and randomizing the order of stimuli for each infant (explained later), we eliminated these possible sources of bias.

Procedure. The infant was seated on the parent's lap facing the yellow center light in the front of the testing booth. The parent listened to instrumental music over headphones to mask the stimuli. In the control room, prior to each trial, the experimenter turned on the yellow center light so that it flashed to draw the infant's attention to the midline.

When the infant fixated on the center light, the experimenter initiated a trial. At this point, one of the side lights began flashing. When the infant looked at the flashing light, speech sounds from either happy BT or neutral ADS categories began playing. During the initial set of familiarization trials, when the infant looked at the side light, it would stop flashing and the sounds would continue as long as the infant looked in the direction of the sound. If the infant looked away for 2 consecutive sec the trial would end, and, after a pause, the experimenter would turn on the center light to attract the infants' attention before beginning another trial.

The infant received a minimum of 15 sec of exposure to each speech style during the familiarization phase. Following this, the test phase began. Trials in the test phase proceeded in similar fashion, except that the side lights continued to flash as long as the infant looked at them. The test phase included 12 trials. Order of presentation (BT vs. ADS) and side of presentation were counterbalanced, and stimuli were presented in random order, selected by computer for each infant, subject to the constraint that there could be a maximum of two consecutive identical trial types (BT, ADS). Multiple utterances were used for each category, and the tokens presented within each category were randomly selected by computer on each trial. Mean looking times were calculated over the six trials for each register during the test phase.

In both phases of the procedure, if the infant failed to orient to the side light within 7.5 sec, the trial would abort and the experimenter would initiate a new trial. If the infant looked at the stimuli for less than 2 sec, the trial would be repeated and only the second measurement would be included in the participant's data. A minimum looking time of 2 sec is required to test sensitivity to suprasegmental features

²Sample stimuli can be downloaded from <http://www.infancyarchives.com>.

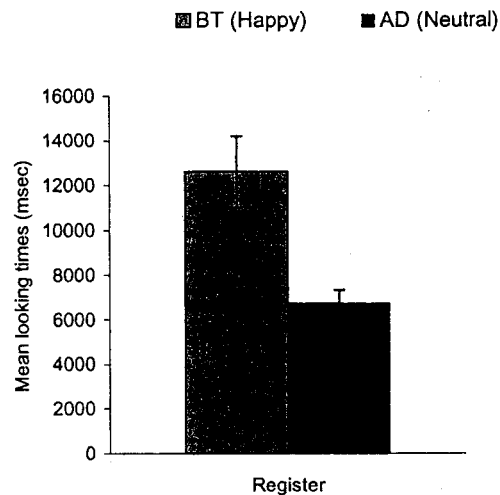


FIGURE 1 Experiment 1: Infants' listening times for happy BT and neutral ADS.

such as affect, which cannot be extracted from an isolable segment. This criterion ensured that infants received sufficient exposure to continuous speech to be able to infer the affective tone of the stimuli. If the infant continued looking at the lights for 30 consecutive sec, the trial would automatically end, and the program would proceed to the next trial.

Results and Discussion

The results of this experiment are shown in Figure 1. As expected, infants showed a significant preference for happy BT over neutral ADS, $t(12) = 3.72, p < .01$. Average looking times were 12.6 sec for BT and 6.7 sec for ADS. Eleven of 13 participants showed a preference for BT. These results, which are consistent with those found in previous experiments, serve to demonstrate the effectiveness of our implementation of the HPP as a suitable procedure for measuring infants' speech preferences and provide prima facie evidence for the validity of our stimulus manipulations.

EXPERIMENT 2

We hypothesized that the BT preferences observed in previous studies may have been due, at least in part, to a persistent confound between speech register and

affect in which BT typically conveyed a happy affective tone, whereas ADS was affectively neutral. If BT preferences are attributable to different affective styles in contrasting stimulus sets, then these preferences might be expected to disappear when affect is held constant across speech registers. Alternatively, if infants exhibit a preference for BT independent of affect, BT preference should persist in different affective contexts. In Experiment 2, BT and ADS stimuli were presented in one of three affective styles in three between-subject stimulus conditions (happy, sad, or neutral). In each case, BT preference was measured to determine if the preference persists when emotion is held constant.

Method

Participants. Thirty-six participants (22 boys and 14 girls) were tested in this study. Data for 4 additional participants were eliminated due to inattention and failure to complete the test. Infants were 6 months old ($M = 179$ days, range = 155–201 days). Recruitment procedures were the same as for Experiment 1.

Stimuli and apparatus. For each infant, the stimuli comprised eight utterances, consisting of four sentences, repeated in ADS and BT registers in happy, sad, or neutral affect. Within each affect condition, infants were exposed to both infant-directed and adult-directed stimuli. For infants in all three conditions, BT stimuli were slower than ADS stimuli (see Table 1). For infants in the happy and sad conditions, BT stimuli were more variable in pitch than ADS stimuli; for infants in the neutral condition, differences in pitch measures across registers were negligible (see Table 2).

Procedure. The same procedure was used in this study as Experiment 1. In this experiment, each infant was randomly assigned to one of three conditions. The experimenter was blind to the assigned condition. The order of presentation (BT or ADS) and side (left or right) of the first test trial was counterbalanced within conditions. Looking times were measured as in the preceding experiment.

Results

The results of Experiment 2 are shown in Figure 2. A mixed affect by register analysis of variance (ANOVA) showed no main effect of register, $F(1, 33) = .404, ns$, no main effect of affect,³ $F(2, 33) = .99, ns$, and no interaction of register and affect,

³In Figure 4, looking times appear to be slightly higher for sad speech than for happy or neutral speech. This slight difference is possibly attributable to the greater length of the sad utterances. If

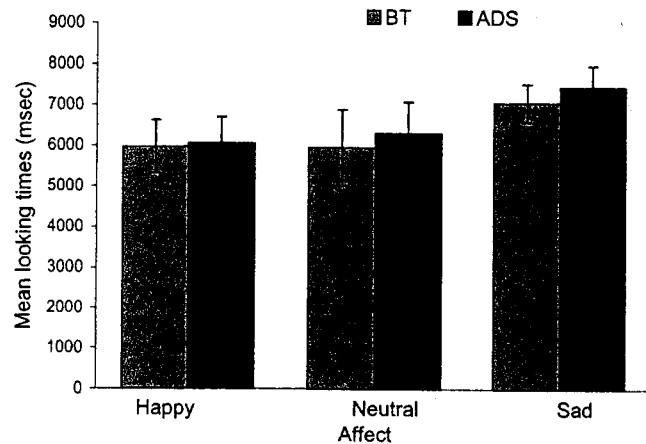


FIGURE 2 Experiment 2: Infants' listening times for register within affect.

$F(2, 33) = .005, ns$. Planned comparisons within each of the three affective conditions failed to reveal any evidence for a BT preference. To the contrary, there appeared to be small but clearly nonsignificant trends toward a preference for ADS in all three conditions. On average, infants hearing happy stimuli listened to ADS for 6 sec and BT for 5.9 sec, $t(11) = 0.09, ns$; 8 infants listened longer to ADS, whereas 4 infants listened longer to BT. In the neutral stimuli condition, infants listened to ADS for 6.3 sec and to BT for 5.9 sec, $t(11) = .54, ns$. Seven infants looked longer at ADS, whereas the remaining 5 looked longer at BT. For sad stimuli, infants looked at ADS for 7.4 sec and to BT for 7.1 sec, $t(11) = .99, ns$. Five infants looked longer at ADS, whereas 7 infants looked longer at BT.

Absolute mean looking times for BT were considerably lower than those observed in Experiment 1. To ensure that our failure to find effects of speech register was not simply due to a quirky sample of inattentive infants, we replicated the happy BT–happy ADS condition with a larger sample in Experiment 2A.

EXPERIMENT 2A

Participants. Twenty-four participants (11 boys and 13 girls) were tested in this study. Data for 1 additional participant were eliminated due to inattention and

infants tended to terminate fixation at the ends rather than the middles of utterances, slightly longer looking times for the longest stimulus set might result. It is important to note that this is a between-subject observation that is not significant. However, this issue will be further explored in Experiment 5, in which sad and neutral speech are presented in a within-subjects design.

failure to complete the test. Infants were 6 months old ($M = 189$ days, range = 173–209 days).

Stimuli, apparatus, and procedure. Infants heard the happy BT and ADS stimuli. Apparatus and procedure were the same as in Experiments 1 and 2, except that in this experiment, the sound was contingent on infants' looking. If an infant looked away from the speaker during a trial the sound would cease immediately. If the infant looked away continuously for 2,000 msec, the trial ended. However, if the infant looked back at the speaker within less than 2,000 msec, the sound would resume.

Results and Discussion

The results of Experiment 2A are shown in Figure 3. As in Experiment 2, infants did not significantly prefer one register to the other, $t(23) = .38, ns$. On average, infants listened to happy BT for 8.5 sec and to happy ADS for 8.3 sec; 14 infants listened longer to BT, whereas 10 infants listened longer to ADS.

Although looking times for BT were somewhat longer than in Experiment 2, they were still considerably shorter than in Experiment 1. This suggests that looking time is partially determined by the contrastiveness of stimulus sets: In Experiments 2 and 2A, infants failed to find the stimulus sets contrastive, and hence their overall looking time was depressed.

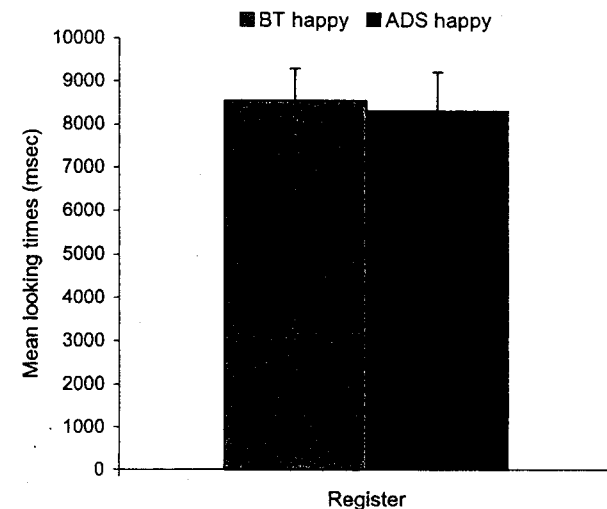


FIGURE 3 Experiment 2A: Infants' listening times for register within happy speech.

The results of Experiments 2 and 2A fail to provide any support for the claim that the BT preference is due to speech register. Nor do they provide any support for claims that the observed preferences are due to higher pitch, more variable pitch, or slower speech rate. Replicating and extending the findings of Kitamura and Burnham (1998), these results hint strongly that the BT preference may be contingent on contrastive affect. When affective valence is controlled, there is no preference for BT over ADS, differences in rated affective intensity between the two registers notwithstanding. The absence of a BT preference extends across both demonstrative and subdued affective contexts. As BT typically incorporates more positive affect than ADS, it follows that in previous reports of a listening preference for BT over ADS, infants may have been attending to positive affect rather than specifically to BT. This suggests that infants may have a fundamental preference for positive emotion in speech.

EXPERIMENT 3

Experiment 3 manipulated affect within registers in two within-subjects conditions. Infants' looking times were measured either for happy versus neutral BT or for happy versus neutral ADS. We expected that results would be consistent with those of Kitamura and Burnham (1998) and Papoušek et al. (1990), in which infants showed evidence of preferences for positive over negative affect. Because the primary goal of this set of studies was to investigate a possible confound introduced by the typical contrast of neutral ADS versus happy ID speech, sad speech was not included in either this experiment or in Experiment 4, although we return to consideration of infant preferences for sad speech in Experiment 5.

Method

Participants. Twenty-eight infants were tested in this procedure (16 boys and 12 girls). Infants were 6 months old ($M = 187$ days, range = 171–202 days). Two infants were excluded due to inattention and failure to complete the test.

Stimuli, apparatus, and procedure. The stimuli used in this study included the happy and neutral tokens used in the previous experiments. Half of the infants heard happy and neutral ADS, whereas the remaining half heard happy and neutral BT. In both conditions, happy stimuli were higher pitched, more variable in pitch, and slower than neutral stimuli (see Tables 1 and 2). Apparatus and procedure were as described in Experiment 1.

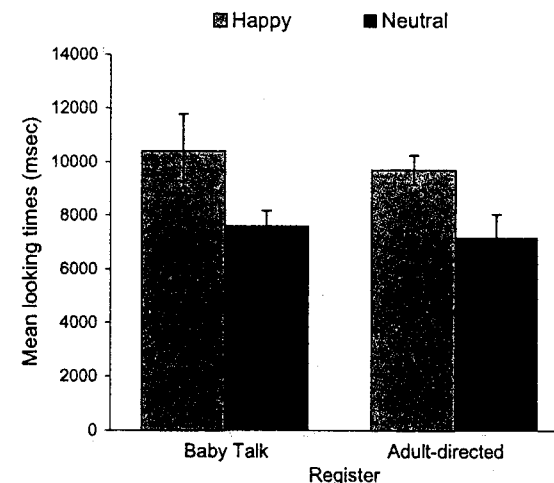


FIGURE 4 Experiment 3: Infants' listening times for affect within register.

Results and Discussion

The results of Experiment 3 are shown in Figure 4. Overall, when speech register was held constant, infants preferred to listen to happy speech more than neutral speech, $F(1, 26) = 11.09, p < .01$. The preference for happy speech was not significantly different between BT and ADS registers, as there was no Affect \times Register interaction, $F(1, 26) = .033, ns$. Eight of 14 infants preferred happy BT to neutral BT; 10 of 14 infants preferred happy ADS to neutral ADS. On average, infants listened to happy BT for 10.2 sec and to neutral BT for 7.6 sec. Average looking times for happy and neutral ADS were 9.7 sec and 7.2 sec, respectively.

These results show that when given a choice between happy and neutral speech, infants are more attracted to positive speech within a given register, and that affect is an important determinant of infants' preferences. This study replicates previous findings of affective preference within BT samples (Kitamura & Burnham, 1998; Papoušek et al., 1990) and extends these findings to ADS samples.

EXPERIMENT 4

Experiments 2 and 2A failed to reveal any preference for BT when affect was held constant, regardless of the type of affect. Experiment 3 revealed a preference for happy talk when register was held constant. If affect is indeed a more powerful mediator of infants' preferences than register so that the BT preference can be

reduced to a general preference for positive emotion, preference for happy ADS over neutral BT as well as for happy BT over neutral ADS would be expected. Experiment 4 tested this prediction.

Methods

Participants. Thirty-two infants were tested in this procedure (18 boys and 14 girls). Infants were 6 months old ($M = 183$ days, range = 171–211 days). One additional infant was excluded from the test due to inattention.

Stimuli, apparatus, and procedure. In this experiment, half of the infants heard happy ADS and neutral BT, whereas the remaining half heard happy BT and neutral ADS. As in Experiment 1, the happy BT stimuli were higher pitched, more variable in pitch, and slower than the neutral ADS stimuli (see Tables 1 and 2). Similarly, the happy ADS stimuli were higher pitched and more variable in pitch than the neutral BT stimuli, although the neutral BT stimuli were slower than the happy ADS stimuli. The apparatus and procedure were the same as in earlier experiments.

Results and Discussion

The results of Experiment 4 are shown in Figure 5. As predicted, when speech register and affect were realigned, the normal pattern of infant preference for BT over ADS was reversed, as shown by the significant crossover interaction, $F(1, 30) = 14.01, p < .001$. Note that this crossover occurred even though our ADS stimuli were rated as less affectively intense than our BT stimuli.

Infants exhibited a significant preference for happy ADS over neutral BT, $t(15) = 2.2, p < .05$, as well as a significant preference for happy BT over neutral ADS, $t(15) = 3.11, p < .01$ (replicating Experiment 1). Eleven of 16 infants preferred happy ADS to neutral BT; mean looking times were 9.5 and 8 sec, respectively. Thirteen of 16 infants preferred happy BT to neutral ADS; mean looking times were 12.9 and 7.8 sec, respectively.

However, the observed crossover interaction appears not to be entirely symmetric. Figure 5 suggests that the magnitude of the preference for happy speech might be greater for happy BT than for happy ADS. To test this possibility, difference scores representing preference for happy speech over neutral speech for a given condition were calculated by subtracting listening times for neutral speech from listening times for happy speech. These difference scores were compared across the two conditions. Infants failed to show a significantly higher difference score in the happy BT and neutral ADS condition compared with the happy ADS and neutral

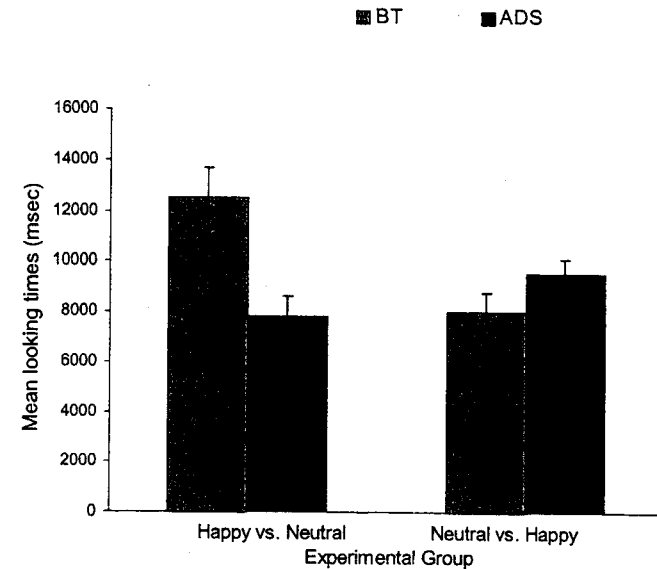


FIGURE 5 Experiment 4: Infants' listening times for happy and neutral BT and ADS.

BT condition, $t(30) = 1.91, p < .10$. Although not significant, this result hints at a possible asymmetry in infants' preferences for happy BT over happy ADS. Figure 5 shows that any difference would have been completely due to the between-subject difference in listening times to happy BT versus happy ADS. However, when we tested this contrast within subjects (in Experiment 2), we found no preference for happy BT over happy ADS.

The results of Experiment 4 might be interpreted as demonstrating that affect plays a predominant role in determining infants' listening preferences. Although these results do show that speech register is not the primary determinant of infants' preferences, there remain additional alternative explanations to be considered. Infants' preferences might be governed by simple prosodic characteristics such as higher pitch, more variable pitch, or increased incidence of expanded intonation contours. In Experiment 2, preferences that might have been expected on any or all of these prosodic characteristics failed to materialize. A stronger demonstration would entail showing that infants' preferences follow affective characteristics of stimuli even when these vary in opposition to these prosodic characteristics.

Moreover, even if it were accepted that the results of this experiment showed that infants' preferences were governed by speech affect, an ambiguity would remain. Infants might prefer emotionally expressive speech, regardless of valence, to emotionally flat speech. Alternatively, infants might have an absolute preference

for happy speech over neutral and sad speech, with no difference between the latter two. Finally, infants might prefer whichever speech is relatively more positive: happy speech over neutral speech or neutral speech over sad speech. The goal of Experiment 5 was to resolve this ambiguity by assessing infants' preferences for neutral versus sad speech.

EXPERIMENT 5

In this experiment, infants' preferences were tested for BT and ADS as in Experiment 4, except that happy stimuli were replaced by sad stimuli. In one condition, infants were tested with sad ADS and neutral BT; in a second condition, infants were tested with neutral ADS and sad BT. If infants are engaged by affective intensity, they might be expected to show a preference for sad speech in both conditions. On the other hand, if infants' preferences are for absolutely positive affect, there should be no difference between sad and neutral stimuli. However, if infants' preferences are guided by the relative positivity of speech affect, it should follow that they would prefer the neutral stimuli in both conditions.

Method

Participants. Twenty-eight infants were tested in this procedure, including 16 boys and 12 girls. Infants were 6 months old ($M = 187$ days, range = 171–230 days).

Stimuli, apparatus, and procedure. The stimuli used in this experiment included the neutral and sad tokens used in the previous experiments. Half of the infants received neutral BT and sad ADS. Gross prosodic differences between the neutral BT and sad ADS stimuli were negligible (the neutral BT stimuli were slightly lower pitched but also slightly less variable in pitch than the sad ADS stimuli; see Table 2). The remaining half of the infants received neutral ADS and sad BT. Table 2 shows that the neutral ADS stimuli were both much lower pitched and much less variable in pitch than were the sad BT stimuli. The apparatus and procedure were the same as in earlier experiments.

Results and Discussion

The results of Experiment 5 are shown in Figure 6. As expected given the results of Experiment 4, infants' listening preferences were not governed by register, as shown by the significant interaction, $F(1, 26) = 6.49, p < .05$. However, rather

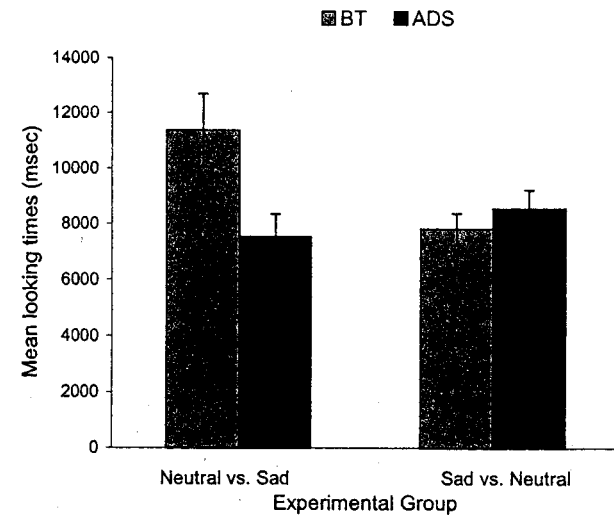


FIGURE 6 Experiment 5: Infants' listening for neutral and sad BT and ADS.

than preferring the more expressive sad speech, infants displayed a significant preference for neutral BT over sad ADS, $t(13) = 2.31, p < .05$. Mean looking times for neutral BT and sad ADS were 11.7 sec and 7.5 sec, respectively. There was a trend toward preference for neutral ADS over sad BT, but it was not significant, $t(13) = 1.07, ns$. Infants listened to neutral ADS for an average of 8.5 sec and to sad BT for an average of 7.8 sec. Nine of 14 infants listened longer to neutral BT than to sad ADS; 8 of 14 infants listened longer to neutral ADS than to sad BT.

Even more clearly than in Experiment 4, the observed interaction appears not to be entirely symmetric: Figure 6 suggests that the magnitude of the preference for neutral speech might have been greater for neutral BT than for neutral ADS. To test this possibility, difference scores representing preference for neutral speech over sad speech for each condition were calculated by subtracting listening times for sad speech from listening times for neutral speech. These difference scores were compared across the two conditions. Infants failed to show significantly higher difference scores in the neutral BT and sad ADS condition compared with the neutral ADS and sad BT condition, $t(26) = 1.73, p < .10$. Figure 6 shows that any difference would have been completely due to the between-subject difference in listening times to neutral BT versus neutral ADS. However, when we tested this contrast within subjects (in Experiment 2), we found no preference for neutral BT over neutral ADS.

These findings show that infants are not drawn to affectively expressive speech independent of valence, nor are they drawn exclusively to affectively positive

speech. Rather, infants' preferences are primarily governed by relative positive affect, independent of register and degree of emotional expressiveness. Clearly, infants do not simply prefer intonational variability to constancy, but rather are drawn to specific acoustic correlates of relative positive emotion. Therefore, infants' preferences represent a functionally meaningful bias, rather than a perceptual bias toward stimulus salience.

GENERAL DISCUSSION

In this research, we sought to determine whether the well-known BT preference might be attributable to the affective properties characteristic of BT rather than to inherent properties that distinguish BT and ADS registers. In a series of experiments that systematically controlled for affect and register, 6-month-old infants' preferences were measured to establish whether the BT preference is contingent on speech affect. Although using spontaneous speech when measuring infants' listening preferences may generally be preferable, Experiment 1 showed that our scripted stimuli, in which MLU, lexical content, and syntactic structure were equated across register, were sufficiently realistic to induce a strong preference for happy BT over neutral ADS. This is precisely the type of contrast that we believe to have been used in previous studies reporting preference for BT over ADS. Experiments 2 and 2A showed that infants' preference for BT over ADS disappears when affect is held constant across speech registers: Of 60 infants tested in these two experiments, exactly half listened longer to each of the two registers. Experiment 3 showed a general preference for happy talk when it was contrasted with neutral talk within each register. If listening preferences are governed primarily by affective content, then it should be possible to reverse the typical patterns of preference by realigning speech register and affect. By pairing positive affect with ADS and neutral affect with BT, Experiment 4 showed that such a reversal could be obtained. Experiment 5 sought to determine whether infants' preferences observed in Experiments 3 and 4 reflected a bias toward affective intensity or affective valence. Results showed that infants are attracted to neutral BT over negative ADS, demonstrating that preference is determined by relative positive affect. Together, these experiments show that the BT preference is contingent primarily on affective components typically found in BT. When those components are incorporated into other registers, infants do not exhibit a preference for BT: Infants' listening preference is not for BT, but rather for happy talk.

Our experiments demonstrate that previous reports of BT preference, which have concluded that the BT register ubiquitously draws infants' attention over the ADS register, have suffered from a persistent confound. These studies did not control for affective stimulus content, which is clearly a powerful mediator of infants' preferences for speech. Previous studies have emphasized the developmental

benefits associated with BT preference. Specifically, it has been claimed that BT educates the infant about native linguistic structure, modulates infants' attention, and draws infants' attention to affective cues (Fernald, 1989, 1992). In light of these findings, we propose a relationship that is the converse of this last claim, whereby it is positive affect that most powerfully draws infants' attention to speech in their environment. To the extent that happy talk reliably predicts BT in infants' environments, it is more likely that an affective bias would direct infants' attention to BT rather than to ADS. If BT presents developmental benefits for language acquisition, it seems natural that it would contain perceptually salient suprasegmental properties, such as positive affect. In addition, affective communication may contribute independently to infants' lexical development, as affectively charged speech contain some of the earliest sound-meaning regularities encountered by young infants (Mastropieri & Turkewitz, 1999).

When faced with speech spoken in contrastive registers, infants' preferences are guided primarily by the degree of positive affect contained in the speech. However, it is highly unlikely that affect represents the sole determinant of infants' listening preferences. Rather, the determinants of attention are likely to be complex and multidimensional. The results of Experiments 4 and 5 hint at the possibility that pitch characteristics—including relative pitch height and variability—may be secondary determinants of infants' listening preferences. In Experiment 5, we observed no preference for neutral ADS versus sad BT. In this instance, although the ADS stimuli possessed more positive affect, the BT stimuli were both higher pitched (by about 100 Hz) and more variable in pitch (by about 9 semitones). This result might therefore be explained by the view that pitch characteristics modulate or temper preferences that are primarily guided by speech affect. This view would suggest that preferences might be strengthened when differences in affect and pitch characteristics are aligned. In Experiment 4, happy BT stimuli possessed more positive affect, higher pitch (by about 60 Hz), and more variable pitch (by about 11 semitones) than neutral ADS stimuli. We observed a trend toward stronger preference in this condition than in the happy ADS versus neutral BT condition, in which the difference in pitch variability across stimulus sets was approximately half an octave smaller. Thus, the slight asymmetries in the interactions in both experiments may be accounted for by appeal to the same mechanism—a secondary influence of pitch characteristics.

It is important to note that pitch characteristics are truly a secondary influence on infants' listening preferences. By themselves, pitch characteristics are not sufficient for determining infants' preferences: In Experiments 2 and 2A, when affect was held constant, no preferences emerged across speech samples that varied in both speech register and pitch characteristics. Our results parallel those of Kitamura and Burnham (1998), who showed that when affect was held constant, no preferences emerged across BT speech samples that varied in pitch characteristics. Further, pitch characteristics are not necessary for determining infants'

preferences. Kitamura and Burnham showed that infants preferred happy BT over neutral BT when pitch characteristics were held constant, and in Experiment 5, we showed that infants preferred neutral BT to sad ADS, even though both pitch height and pitch variation were equivalent across the two samples. Other factors typically distinguishing BT and ADS, such as MLU, lexical content, or syntactic structure, all of which were controlled in our stimuli, may also serve as secondary influences on infants' listening preferences.

The primacy of affect in determining listening preference mirrors the development of productive language. Infants use intonation to express desires and needs long before they begin to express referential meanings with words (Lewis, 1936), displaying an early sensitivity to the power of intonation as a communicator of affect. Articulatory differences aside, this is perhaps due to the fact that infants can exploit affective cues from birth, whereas language must be learned (Bloom, 1990). In addition, the cognitive requirements for affect communication and language development are different. Affective expression requires appraisal of an internal state, which can be communicated nonlinguistically at a very early age. In contrast, words are arbitrary units, so language comprehension requires the formation of mental representations and memory for lexical items, which requires cognitive effort on the part of the language learner (Bloom, 1990). Therefore, infants can use paralinguistic means to communicate information about their internal state from birth, although they cannot use linguistic means to communicate affective states until much later in development. Interestingly, newborns only show differential responsiveness to positive and negative vocal affect when stimuli are spoken in their native language (Mastropieri & Turkewitz, 1999). This suggests that vocal affect perception may develop prenatally based on associations between prosodic characteristics of maternal affective expression and the physiological changes that accompany these vocalizations.

These findings complement those of Trainor et al. (2000) on the role of affect in adults' productions to infants. Their analyses of BT and ADS revealed few acoustic differences between registers when affect was controlled, from which they concluded that the production of BT arises from vocal affect expression. This suggests that adults are more inclined to manipulate prosodic features of their speech and communicate affectively with infants, perhaps because infants are attuned to vocal emotion in their environment as a marker for speech directed to them. Further, it has been shown that adults rate infants' facial responses to BT as more attractive than their facial responses to ADS (Werker & McLeod, 1989). To the extent that speech to infants contains positive affect, infants' facial responses to positive affect may serve to facilitate parent-child interaction.

The antecedence of affective expression over linguistic communication in evolution was documented by Darwin (1872) in his extensive discussion of the similarities in affective expression in nonhuman and human primates. Darwin discussed the necessity of vocal and facial affect expression in determining communicative intent

and internal state information among conspecifics in primate communities to distinguish between sources of threat and allegiance. The affective properties of BT are well suited to the neurological immaturity of infants' auditory systems, which require an initial dependence on subcortical auditory structures. These structures effect holistic analyses of speech, such as the extraction of affective content. It is logical to reason that developing systems will process speech using phylogenetically older and simpler mechanisms and later proceed to more complex linguistic analyses (Fernald, 1992). Fernald (1992) elaborated on the long-term adaptive advantage of early emotional communication in instilling an important element of social competence, which enhances reproductive fitness. Certainly, there are reported similarities in human infants' and nonhuman primates' abilities to discriminate prosody across languages (Ramus, Hauser, Miller, Morris, & Mehler, 2000).

Given the evolutionary primacy of affect, it seems reasonable that affective cues should guide infants' early preferences for speech. An early preference for positive affect has also been documented in the visual domain, where an innately specified tendency to attend to positive affect has been proposed (Serrano et al., 1992), given the functional significance of happy faces. These findings corroborate an early preference for positive affect expression in the environment in a different sensory modality. It remains unclear whether preferences for happy faces and voices emerge concurrently or whether affective sensitivity in one modality fuels that in other modalities. Although neither facial nor vocal affect may be intrinsically reinforcing in the absence of experience, the perception of vocal affect may represent some of the most accessible sound-to-meaning correspondences that infants can extract from the input.

The functional significance of attending to happy talk has been considered by Scherer (1986) in his review of vocal affect communication. He proposed that positive affect in speech predicts an approach reaction in humans, whereby it captures their attention more than neutral or negative speech. Scherer's review suggests that this general preference for happiness in speech is conducive to human development, in that it encourages infants to pay attention to familiar figures, such as caregivers, and other nonthreatening sources of support, which is particularly important in altricial species. In contrast, negative speech elicits a withdrawal reaction as a negative emotional state poses a threat to the listener, predicting that the organism will likely benefit from selectively attending to positive vocal affect.

With respect to this study, positive vocal affect may serve as an intrinsically pleasant stimulus, drawing the infants' attention to happy talk, which includes BT. From an adaptive perspective, attending to BT may facilitate inculcation of native language patterns, an important developmental goal. Affect may serve as a very early instrument to direct the infant to BT to galvanize a curiosity for the relationship between native sound patterns and linguistic units in the BT speech stream. Although previous investigations of infants' sensitivity to vocal

affect have tested infants younger than those here (e.g., Fernald, 1993; Kitamura & Burnham, 1998; Papoušek et al., 1990), it is possible that later in development affective cues are isolated from register and perceived as independent communicators of speaker intent. Therefore, BT preference may be a dynamic process, initially governed by affective prosody and later driven by attention to language-specific aspects of the input that appeal to infants. This is consistent with previous reports that parent–infant communication may change from an affectively based register to a referentially based register over the first year of life (Werker & McLeod, 1989). This issue could be addressed by investigating whether the bases of newborns' preferences for BT are primarily acoustic or functional. Similarly, adults may adapt their speech to accommodate a shift in focus from language-general to native features of the input.

We speculate that production of and preference for happy talk is of evolutionarily determined benefit for the regulation and maintenance of infant–caretaker relations. However, it remains unclear what processing advantages there are for happy talk. Research by Golinkoff and Alioto (1995) suggested that there might be such advantages to BT, showing that adults more readily acquire novel words when they are spoken in BT than ADS. However, no comparable demonstration with infants has yet been forthcoming. Inasmuch as preference indexes the conscious allocation of attention, we venture to hypothesize that happy talk will have varying effects on tasks involved in acquisition, depending on the degree to which controlled processing is required. For example, the recognition of words in fluent speech engages automatic processes; on these grounds we might expect that happy talk would have limited impact on this task. Indeed, a set of follow-up studies now in progress examines whether the inclusion of positive affect in speech enhances word recognition in preverbal infants. Preliminary results from these studies indicate that although affective matching between familiarization and recognition stimuli is important for younger infants, neither the strength nor valence of the affect has significant import.

At the other end of the spectrum, word selection in speech production is highly controlled. If we consider the first words that infants select to produce, it is easy to see that they generally have pleasant associations and are likely to have been encountered most often in happy talk. The task of learning word meanings falls somewhere between word recognition and word selection; therefore, we would expect that tasks involving inculcation of word–meaning correspondences would display moderately beneficial effects of happy talk.

This set of studies demonstrates happy talk is the primary factor underlying the BT preference. Our findings reveal a confound in reports of BT preference that have failed to control for affect, thereby challenging the conclusion that BT uniquely invites infants' early analyses of the speech stream. Positive vocal affect attracts infants' attention independent of the register into which it is incorporated; thus, previous demonstrations of BT preference are best reinterpreted as revealing a more general preference for happy talk. That infants' attention should be elicited by

positive affect makes perfect evolutionary sense: Selective attention to happy vocalizations is as adaptive for human infants as it is for nonhuman primates. Both human and nonhuman primates use affect to communicate intention or desires to their young (Darwin, 1872). While humans can be distinguished from nonhuman primates in the complexity and sophistication of their communication system, the origins of symbolic communication are likely to be founded on common evolutionary primitives such as affect (Fernald, 1992). The instinctive inclusion of positive emotion in speech to infants may help to ensure that infants attend to optimal input, providing an example of one of perhaps myriad ways in which language, as it has developed, has opportunistically capitalized on preexisting capacities.

ACKNOWLEDGMENTS

This research was supported by a grant from the National Institute of Health (5 R01 HD32005) to James L. Morgan. We are grateful to Nancy Allard, Karen Rathbun, and Crystal Wilson for their assistance with data collection and stimulus development. We also thank Heather Bortfeld, Robin Cooper, Katherine Demuth, Wendy Ostroff, Julie Sedivy, and Katherine White for comments on drafts.

REFERENCES

- Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of Personality and Social Psychology, 70*, 614–636.
- Bernstein-Ratner, N., & Pye, C. (1984). Higher pitch in BT is not universal: Acoustic evidence from Quiche Mayan. *Journal of Child Language, 11*, 515–522.
- Bloom, L. (1990). Developments in expression: Affect and speech. In N. L. Stein, B. Leventhal, & T. Trabasso (Eds.), *Psychological and biological approaches to emotion* (pp. 215–246). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Brown, R. (1977). Introduction. In C. Snow & C. Ferguson (Eds.), *Talking to children: Language input and acquisition* (pp. 1–30). Cambridge, England: Cambridge University Press.
- Cooper, R. P., & Aslin, R. N. (1990). Preference for infant-directed speech in the first month after birth. *Child Development, 61*, 1584–1595.
- Cooper, R. P., & Aslin, R. N. (1994). Developmental differences in infant attention to the spectral properties of infant-directed speech. *Child Development, 65*, 1663–1677.
- Cruttenden, A. (1994). Phonetic and prosodic aspects of baby talk. In C. Gallaway & B. J. Richards (Eds.), *Input and interaction in language acquisition* (pp. 135–152). Cambridge, England: Cambridge University Press.
- Darwin, C. (1872). *The expression of the emotions in man and animals*. London: Murray.
- DeCasper, A. J., & Fifer, W. P. (1987). Of human bonding: Newborns prefer their mothers' voices. In J. Oates & S. Sheldon (Eds.), *Cognitive development in infancy* (pp. 111–118). Hove, England: Lawrence Erlbaum Associates, Inc.
- Ekman, P., & Friesen, W. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology, 17*, 124–129.

- Fee, J., & Shaw, K. (1998). Pitch modifications in Mi'kmaq child-directed speech. In E. Clark (Ed.), *The proceedings for the 29th annual child language research forum* (pp. 47–54). Chicago: Center for the Study of Language and Information.
- Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development*, 8, 181–195.
- Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development*, 60, 1497–1510.
- Fernald, A. (1992). Human maternal vocalizations to infants as biologically relevant signals: An evolutionary perspective. In J. H. Barkow, L. Cosmides, & J. Toobey (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture* (pp. 391–428). New York: Oxford University Press.
- Fernald, A. (1993). Approval and disapproval: Infant responsiveness to vocal affect in familiar and unfamiliar languages. *Child Development*, 64, 657–674.
- Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development*, 10, 279–293.
- Fernald, A., & Simon, T. (1984). Expanded intonation contours in infants' speech to newborns. *Developmental Psychology*, 20, 104–113.
- Fernald, A., Taeschner, T., Dunn, J., Papoušek, M., Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modification in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, 16, 477–501.
- Fisher, C., & Tokura, H. (1996). Acoustic cues to grammatical structure in infant-directed speech: Cross-linguistic evidence. *Child Development*, 67, 3192–3218.
- Furrow, D., Nelson, K., & Benedict, H. (1979). Mothers' speech to children and syntactic development: Some simple relationships. *Journal of Child Language*, 6, 423–442.
- Garnica, O. K. (1977). Some prosodic and paralinguistic features of speech to young children. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children: Language input and acquisition* (pp. 63–88). Cambridge, England: Cambridge University Press.
- Gleitman, L. R., Newport, E. L., & Gleitman, H. (1984). The current status of the motherese hypothesis. *Journal of Child Language*, 11, 43–79.
- Golinkoff, R., & Alioto, A. (1995). Infant-directed speech facilitates lexical learning in adults hearing Chinese: Implications for language acquisition. *Journal of Child Language*, 22, 703–726.
- Kappas, A., Hess, U., & Scherer, K. (1991). Voice and emotion. In R. Feldman & E. Rime (Eds.), *Fundamentals of nonverbal behavior* (pp. 200–238). New York: Cambridge University Press.
- Kemler Nelson, D., Hirsh-Pasek, K., Jusczyk, P. W., & Wright-Cassidy, K. (1989). How the prosodic cues in motherese might assist language learning. *Journal of Child Language*, 16, 55–68.
- Kemler Nelson, D. G., Jusczyk, P. W., Mandel, D. R., Myers, J., Turk, A., & Gerken, L. A. (1995). The headturn preference procedure for testing auditory perception. *Infant Behavior and Development*, 18, 111–116.
- Kitamura, C., & Burnham, D. (1998). The infant's response to maternal vocal affect. In C. Rovee-Collier, L. Lipsitt, & H. Hayne (Eds.), *Advances in infancy research* (Vol. 12, pp. 221–236). Stamford, CT: Ablex.
- Lewis, M. M. (1936). *Infant speech: A study of the beginnings of language*. London: Routledge & Kegan Paul.
- Mastropieri, D., & Turkewitz, G. (1999). Prenatal experience and neonatal responsiveness to vocal expressions of emotion. *Developmental Psychobiology*, 35, 204–214.
- McLeod, P. J. (1993). What studies of communication with infants tell us about psychology—Baby talk and other speech registers. *Canadian Psychology*, 34, 282–292.
- McRoberts, G. W., & Whaley, S. (1995, March). *Five-month-old infants prefer ID speech in Japanese and Xhosa*. Poster presented at the biennial meeting of the Society for Research in Child Development, Indianapolis, IN.
- Mehler, J., Dupoux, E., Nazzi, T., & Dahan-Lambertz, G. (1996). Coping with linguistic diversity: The infant's viewpoint. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax* (pp. 101–116). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Moon, C., Cooper, R. P., & Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant Behavior and Development*, 16, 495–500.
- Morgan, J. L. (1986). *From simple input to complex grammar*. Cambridge, MA: MIT Press.
- Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 756–766.
- Papoušek, M., Bornstein, M. H., Nuzzo, C., Papoušek, H., & Symmes, D. (1990). Infant responses to prototypical melodic contours in parental speech. *Infant Behavior and Development*, 13, 539–545.
- Pegg, J. E., Werker, J. F., & McLeod, P. J. (1992). Preference for infant-directed over adult-directed speech: Evidence from seven-week-old infants. *Infant Behavior and Development*, 15, 325–345.
- Pinto, J. P., Fernald, A., McRoberts, G. W., & Cole, S. (1999). Reliability and validity in infant auditory preference procedures. In C. Rovee-Collier, L. Lipsitt, & H. Hayne (Eds.), *Advances in infancy research* (Vol. 12, pp. 221–236). Stamford, CT: Ablex.
- Ramus, F., Hauser, M. D., Miller, C., Morris, D., & Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science*, 288, 349–351.
- Sachs, J. (1977). The adaptive significance of linguistic input to prelinguistic infants. In C. Snow & C. Ferguson (Eds.), *Talking to children: Language input and acquisition* (pp. 32–50). Cambridge, England: Cambridge University Press.
- Serrano, J. M., Iglesias, J., & Loeches, A. (1992). Visual-discrimination and recognition of facial expressions of anger, fear, and surprise in 4-month-old to 6-month-old infants. *Developmental Psychobiology*, 25, 411–425.
- Serrano, J. M., Iglesias, J., & Loeches, A. (1995). Infants' responses to adult static facial expressions. *Infant Behavior and Development*, 18, 477–482.
- Scherer, K. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin*, 9, 143–165.
- Snow, C. E. (1972). Mothers' speech to children learning language. *Child Development*, 43, 549–565.
- Snow, C. E. (1999). Social perspectives on the emergence of language. In B. MacWhinney (Ed.), *The emergence of language* (pp. 257–276). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Tartter, V., & Braun, D. (1994). Hearing smiles and frowns in normal and whisper registers. *Journal of Society of America*, 96, 2101–2107.
- Trainor, L. J., Austin, C. M., & Desjardins, R. N. (2000). Is infant-directed speech prosody a result of the vocal expression of emotion? *Psychological Science*, 11, 188–195.
- Werker, J. F., & McLeod, P. J. (1989). Infant preference for both male and female infant-directed talk: A developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology*, 43, 230–246.

APPENDIX: SAMPLE PASSAGES USED IN STIMULUS ELICITATION

Happy: I hadn't seen Grandma in ages. She lived 200 miles away and traveled regularly. She had been unable to visit us for a while. I hoped that we would be able to see her soon as Christmas was approaching. I really loved spending time with her. Just as I was remembering her fondly, the most incredible thing happened.

Dad entered the room. He told us all to get ready and packed for the weekend. "We're going to Grandma's house!" he announced.

Neutral: I was watching TV when my mom came in. She then asked me to finish my homework tonight. I asked her why I couldn't do it later. "We're going to Grandma's house," she replied. I agreed and she left the room.

Sad: I used to dread Christmas. We would always go to Grandma's house. I would get bored beyond belief. I never enjoyed staying in her home. She was eternally dissatisfied with life. It was always a depressing affair. Tomorrow was Christmas Eve and I was not looking forward to it. I even hate to think about it. We're going to Grandma's house. I wish this holiday were over.

Happy: Today is my birthday. My mom promised me that we would have my favorite dinner. I was hoping that she would make spaghetti. I would be so happy if we were having spaghetti; it would make my day! Finally, she called us down for dinner. As I entered the kitchen, my face lit up. "We're having spaghetti for dinner!" I exclaimed. I was so happy!

Neutral: I was setting the table for dinner. My mom was cooking and my dad was reading outside. We're having spaghetti for dinner. After that, I'm going to do my homework and then go to sleep.

Sad: It had been a really tough day. I was really tired and I had failed an exam. I felt despondent as I walked into the cafeteria. I was hoping that dinner would be good. Anything but spaghetti would be fine. I would be really disappointed if we were having spaghetti. It would be another terrible reason why this day is insufferable. As I approached the food, I was really disappointed. We're having spaghetti for dinner. I couldn't wait for this day to be over.