Infant Auditory Sensitivity to Pure Tones and Frequency-Modulated Tones

Lori J. Leibold
Department of Allied Health Sciences
University of North Carolina at Chapel Hill

Lynne A. Werner
Department of Speech and Hearing Sciences
University of Washington

It has been suggested that infants respond preferentially to infant-directed speech because their auditory sensitivity to sounds with extensive frequency modulation (FM) is better than their sensitivity to less modulated sounds. In this experiment, auditory thresholds for FM tones and for unmodulated, or pure, tones in a background of noise were measured for 4-month-old infants using a conditioned response procedure. The FM tones swept from 150 to 275 Hz or from 150 to 550 Hz. The frequency of the pure tone was either 275 or 550 Hz. The results showed that infants were slightly, but significantly, more sensitive to the sounds that included 550 Hz than they were to the lower frequency sounds, whether or not the sound was frequency modulated. It appears that infants could be somewhat more sensitive to infant-directed than to adult-directed speech, not because of FM per se, but because the fundamental frequency excursions in infant-directed speech extend into a higher frequency range than those in adult-directed speech.

When adults talk to infants, they tend to exaggerate intonation compared to speech directed to another adult (e.g., Fernald & Simon, 1984). Moreover, infants will listen longer to such infant-directed (ID) speech than to adult-directed (AD) speech (e.g., Fernald, 1985). ID speech shares many acoustic characteristics with speech directed to hearing-impaired adults, to normal-hearing adults in noisy environments, and to elderly and foreign speakers (e.g., Bradlow & Bent, 2002; Kemper, Ferrell, Harden, Finter-Urczyk, & Billington, 1998; Krause & Braida, 2004;

Picheny, Durlach, & Braida, 1986; Liu & Zeng, 2006). These include slower speaking rate, greater temporal modulation and fundamental frequency variance, expanded vowel space, and greater emphasis on higher frequencies (Krause & Braida, 2002, 2004; Liu, Del Rio, Bradlow, & Zeng, 2004; Payton, Uchanski, & Braida, 1994; Picheny et al., 1986; Uchanski, Choi, Braida, Reed, & Durlach, 1996; Zeng & Liu, 2006). These characteristics make speech more audible, more perceptually segregated from background noise, and more intelligible (e.g., McAdams, 1989; Picheny, Durlach, & Braida, 1985). Thus, ID speech may not represent a unique communication style, but rather adults' attempt to compensate for immature auditory sensitivity (Olsho, Koch, Carter, Halpin, & Spetner, 1988; Trehub, Schneider, & Endman, 1980) or for infants' greater susceptibility to interference from background sounds (e.g., Werner & Bargones, 1991).

Adults may adopt an ID speech register, then, because infants are more sensitive to ID speech than to AD speech. Typically, better sensitivity means a lower detection threshold. Lower detection thresholds are associated with some auditory advantages at suprathreshold levels. For example, if two similar moderate level sounds with different detection thresholds are presented at the same sound pressure level, the sound with the lower detection threshold will sound louder. Likewise infants only achieve adult levels of speech discrimination performance when speech intensity is increased to compensate for their poorer detection thresholds (Nozza, Rossman, Bond, & Miller, 1990; Nozza, Wagner, & Crandell, 1988). More important, differences in detection threshold can indicate differences in suprathreshold perceptual processes. For example, the ability to perceptually segregate speech from competing sounds leads to lower thresholds for the speech as well as improved intelligibility at suprathreshold levels. Auditory advantages in the latter cases are often as great as 10 to 15 dB (e.g., Beutelmann & Brand, 2006).

Colombo, Frick, Ryther, Coldren, and Mitchell (1995) concluded that infants are more sensitive to sounds with the frequency contours of ID speech, because infants responded to the addition of a frequency-modulated (FM) tone analog of ID speech to a noise, whereas they did not respond to the addition of either an FM tone analog of AD speech or a pure tone to the same noise. The habituation and recovery paradigm used by Colombo et al., however, is not well suited to measuring thresholds, and the possibility that infants responded differentially to the ID FM tone because they found it more interesting cannot be eliminated.

Newman and her colleagues have completed several studies related to this issue. They have found that, against a background of AD speech, adults are better able to follow a speaker with an ID style than a speaker with an AD style (Newman, Weppelman, & Hussain, 2003). This supports the idea that ID speech promotes segregation from background sounds. Further, although mothers speaking to toddlers use an ID speaking style in quiet, they make some additional modifications to their speech when noise is introduced (Newman, 2003). This finding supports the idea that parents adjust the way they talk to their children with

changes in the communication context. However, Newman and Hussain (2006) showed recently that 4.5-month-old infants do not show greater preference for ID speech in a noisy background than they do in quiet. This finding argues against the idea that infants prefer ID speech because they can hear it better, but as Newman and Hussain pointed out, infants may still hear ID speech better but exhibit no preference for it over AD speech.

The primary goal of this experiment was to determine whether infants' detection thresholds for an FM tone analog of ID speech are lower than their thresholds for an FM tone analog of AD speech. A method developed specifically to test infants' auditory sensitivity (Werner, 1995) was used. The FM tones were identical to the 150 to 275 Hz and 150 to 550 Hz sweeps used by Colombo et al. (1995). Pure tones of 275 and 550 Hz were used as unmodulated comparison signals. Including the pure tones made it possible to determine whether it is the presence of FM or the presence of higher frequencies in the ID FM sweep that is responsible for any threshold differences. If FM helps infants to perceptually segregate the tone from the background noise, the difference between thresholds for pure tones and FM tones could be on the order of 10 dB. However, to adults, a 550-Hz tone will be a few dB more audible than a 275-Hz tone, because masked threshold tends to grow relatively worse with decreasing frequency below about 500 Hz (Patterson & Moore, 1986).[1]

METHOD

Participants

Fifty-seven infants participated, with an average age at the initial testing session of 17.4 weeks (SD = 1.1 week).

Infants had no risk factors for hearing loss as reported by a parent, had suffered two or fewer episodes of otitis media and no episode within the prior week, and were healthy on the test day. All infants passed screening for middle ear fluid on the test day. The data from an additional 72 infants were excluded from analysis: 33 infants did not reach training criterion; 20 infants did not provide sufficient test data; and 19 infants completed testing but were excluded because of high variability of reversals, low response rate on probe trials, or high false alarm rate (see later). Examination of the thresholds of infants who reached training criteria, but whose data did not meet other inclusion criteria, showed that the excluded thresholds were slightly higher and more variable than those included in the analyses but were similarly affected by stimulus condition.

---

[1]Masked threshold increases again as frequency increases above 500 Hz (Zwicker & Fastl, 1990).

[2]If thresholds were included from infants who completed testing, but whose data did not meet the inclusion criterion, the frequency effect was marginally significant, p = .08.

Stimuli and Procedure

The stimuli were FM tones, pure tones, and a broadband noise masker. All of the stimuli matched those used by Colombo et al. (1995) except for the 550 Hz tone, which was not used in that experiment. The FM tones mimicked the frequency contour of AD speech (150–275 Hz FM tone) or ID speech (150–550 Hz FM tone). Pure tones of 275 and 550 Hz were the highest frequency in each FM tone. All signals were 1 sec in duration, including a 16-msec rise and fall time. The FM tones swept linearly from the lowest to the highest and back to the lowest frequency in 1 sec. Signals were presented in a continuous background of broadband noise. The spectrum level of the noise was 25 dB SPL. Stimuli were digitally generated and low-pass filtered at 4,000 Hz using TDT system III programmable hardware. Custom software was used to control the experiment. Spectrum analysis confirmed the frequency content of the signals; the spectrum of the noise was flat below 1,000 Hz.

Each infant was randomly assigned to participate in one of the four signal conditions. Infants were tested in two visits within 2 weeks, each lasting approximately 45 min.

Infants were tested in a single-walled, sound-attenuating room using an observer-based procedure (Werner, 1995). Infants sat on their parent's lap in the room throughout testing. The stimuli were delivered to the infant's right ear through an ER1 insert earphone, through a foam tip. To keep the infant facing toward midline, an assistant sat in the room to the left of the parent and infant and manipulated quiet toys. Masking sounds were presented to both the parent and the assistant through circumaural headphones. To the right of the infant and parent were two mechanical toys enclosed with lights in a dark Plexiglas box. The noise was presented continuously in the background throughout the session.

An observer sat outside the booth and watched the infant through a window and on a video monitor. The observer began a trial when the infant was quiet and facing ahead. On each trial either a signal (FM tone or pure tone) or a no signal (noise only) was presented. The observer did not know the trial type, but judged whether a signal or no signal had been presented on each trial, based on the infant's behavior. The observer received feedback after every trial. The infants learned to respond when they heard a signal because the mechanical toys were illuminated and activated as reinforcement whenever the observer correctly identified a signal trial. Typical infant responses included changes in motor activity and looking in the direction of the mechanical toys.

Sessions began with two training phases. The signal level was expected to be clearly audible in both training phases, based on previous studies and on pilot data. The purpose of the first training phase was to demonstrate the relation between the reinforcer and the signal. The purpose of the second training phase was to teach the

infant that a response was required on a signal trial to turn on the toy. Training ended when the observer achieved a hit rate of at least .8 and a false alarm (signal response on a no-signal trial) rate of no greater than .2. The average number of trials required to complete both training phases was 27.5 ($SD = 11.5$), with no differences across conditions.

In the testing phase, detection thresholds were determined adaptively: Signal level decreased if the signal was correctly detected on two consecutive trials and increased if the signal was missed on any trial (Levitt, 1971). The starting level of the signal was approximately 10 dB higher than the expected threshold. The initial step size was 6 dB and varied during testing following Parameter Estimation by Sequential Tracking (PEST) rules (Taylor & Creelman, 1967). Testing ended when eight reversals were obtained and threshold was calculated as the average of the last six reversals. The probability of a signal trial was .75 and the probability of a no-signal trial was .125. In addition, probe signals at the training level were presented with a probability of .125. Thresholds were accepted only if the false alarm response rate was lower than .4, the probe response rate was greater than .6, and eight reversals with the last six within a 10 dB range were obtained. The number of thresholds obtained in each of the four conditions ranged from 13 to 16.

RESULTS

Average thresholds for each condition are shown in Figure 1. Thresholds were lowest for the 150 to 550 Hz FM tone, followed by the 550 Hz pure tone, the 150 to 275 Hz FM tone, and the 275 Hz pure tone. Thus, thresholds for the sounds containing the higher frequencies appeared to be a few dB lower than thresholds for
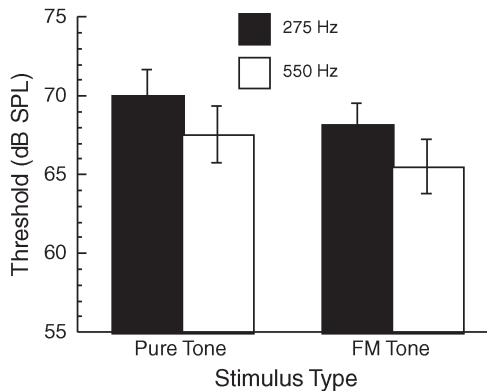

FIGURE 1   Average threshold as a function of stimulus type for stimuli with a maximum frequency of 550 or 275 Hz. Error bars represent ±1 *SEM*.

the lower frequencies; thresholds for the FM tones appeared to be slightly lower than thresholds for the pure tones.

A frequency (high or low) × stimulus type (FM or pure tone) analysis of variance (ANOVA) was performed. The interaction term was not significant, $F(1, 53) = .04$, $p = .84$, $\eta_p^2 = .001$. Thresholds for the 150 to 550 Hz FM tone and the 550 Hz pure tone were significantly lower than thresholds for the 150 to 275 Hz FM tone and the 275 Hz pure tone, $F(1, 53) = 5.88$, $p = .02$, $\eta_p^2 = .100$. However, the FM tone thresholds were not significantly different from the pure tone thresholds, $F(1, 53) = 1.51$, $p = .23$, $\eta_p^2 = .028$.

Two additional analyses were completed to ensure that the effects reported were not the result of across-condition differences in response bias or in asymptotic level of performance. Both of these variables can influence the accuracy with which an adaptive technique estimates threshold. First, Frequency × Stimulus Type ANOVAs of false alarm rate, of probe response rate, and of the range of the psychometric function (the difference between probe response rate and false alarm rate) indicated no significant main effects or interaction (all ps > .11, $\eta_p^2 < .05$). Thus, there was no evidence that these variables differed across stimulus conditions. The average false alarm rate was .31 ($SD = .12$), and the average probe hit rate was .88 ($SD = .15$).

Second, a psychometric function, p(hit) as a function of level, was fit to the pooled data of all infants in each condition using probit analysis (Finney, 1970). Responses were combined over a range of 3 to 4 dB. The observed false alarm rate and hit rate for intensities over 80 dB were specified as the lower and upper asymptotes of the functions, respectively. The resulting functions range from 0 to 1, because they are effectively rescaled based on the false alarm rate and the asymptotic hit rate. They represent the predicted sensitivity of the infants, independent of false alarm rate and asymptotic hit rate. Thresholds calculated from pooled psychometric functions would be expected to produce thresholds and threshold differences similar to those obtained from the adaptive procedure, although the slopes may not reflect the slopes of the individual infants' psychometric functions accurately. The best fitting psychometric functions are shown in Figure 2. Note that the functions for the 150 to 550 Hz FM tone and the 550 Hz pure tone fall to the left of the functions for the 150 to 275 Hz tone and the 275 Hz pure tone (i.e., at lower intensities). Thresholds were estimated from the fitted pooled psychometric functions as the level yielding a 0.7 p(hit), the value on which the adaptive procedure should have converged. The thresholds differed from those plotted in Figure 1 by 1 to 3 dB, but the pattern of results paralleled that seen in the adaptive thresholds. The difference between the 550 Hz and 275 Hz pure tone thresholds was 3.6 dB, and the difference between the 150 to 550 Hz and 150 to 275 Hz FM tone thresholds was 4.9 dB. The differences between the pure tones and FM tones were 1 to 2 dB. Thus, the results were generally consistent across analyses.
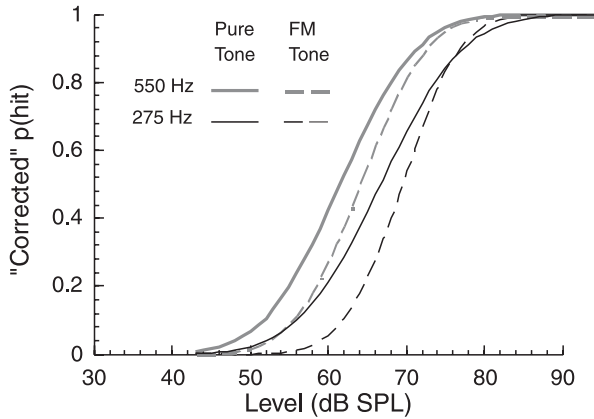
FIGURE 2 Best fitting pooled psychometric functions for FM tones (dashed lines) and pure tones (solid lines) with maximum frequency of 550 Hz (thick gray lines) and 275 Hz (thin black lines). The functions were fitted using probit analysis with the observed false alarm rate and asymptotic hit rate as parameters. Corrected *p* (hit) refers to the fact that the observed *p*(hit) is effectively adjusted for false alarm rate and asymptotic hit rate in the fitting process.

DISCUSSION

The results of this study indicate that infants have slightly, but significantly, lower thresholds for an FM tone that mimics the frequency contour of ID speech than they do for an FM tone that mimics the frequency contour of AD speech. The same threshold difference is seen between two pure tones with frequencies matching the highest frequencies in the ID and AD FM sweeps, respectively. Similar frequency effects in pure tone detection in noise are reported in published studies of adult hearing (e.g., Moore, Peters, & Glasberg, 1993). The results reported here suggest that the difference between FM tones in detection threshold is a consequence of the fact that the highest frequencies in the ID sweep are detectable at lower intensities than the highest frequencies in the AD sweep. The thresholds obtained here are consistent with the results of Colombo et al. (1995), because the sound levels used in that study would have been just above the threshold reported here for the 150 to 550 Hz FM sweep, and just at or below the thresholds reported here for the 150 to 275 Hz FM sweep and the 275 Hz pure tone.

These results indicate that infants may be somewhat more sensitive to ID speech than to AD speech, because ID speech contains higher frequencies than AD speech. One unanswered question is whether or not the threshold difference observed for FM tones can be generalized to ID and AD speech. Fundamental frequency excursion, independent of the frequency range, may be more important for

the perceptual segregation of speech from noise than of a tone from noise. If that were the case, the perceptual advantage of ID speech over AD speech may be greater than that observed for tones. Another issue not directly addressed here is the role that sensitivity may play in infants' preference for ID speech over AD speech. Although better sensitivity to ID speech than to AD speech, especially of the magnitude observed here, may not be solely responsible infants' preference for ID speech, it may still contribute to that preference.

REFERENCES

Beutelmann, R., & Brand, T. (2006). Prediction of speech intelligibility in spatial noise and reverberation for normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America, 120,* 331–342.

Bradlow, A. R., & Bent, T. (2002). The clear speech effect for non-native listeners. *Journal of the Acoustical Society of America, 112,* 272–284.

Colombo, J., Frick, J. E., Ryther, J. S., Coldren, J. T., & Mitchell, D. W. (1995). Infants' detection of analogs of "motherese" in noise. *Merrill-Palmer Quarterly, 41,* 104–113.

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development, 8,* 181–195.

Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology, 20,* 104–113.

Finney, D. J. (1970). *Probit analysis.* Cambridge, UK: Cambridge University Press.

Kemper, S., Ferrell, P., Harden, T., Finter-Urczyk, A., & Billington, C. (1998). Use of elderspeak by young and older adults to impaired and unimpaired listeners. *Aging Neuropsychology and Cognition, 5,* 43–55.

Krause, J. C., & Braida, L. D. (2002). Investigating alternative forms of clear speech: The effects of speaking rate and speaking mode on intelligibility. *Journal of the Acoustical Society of America, 112,* 2165–2172.

Krause, J. C., & Braida, L. D. (2004). Acoustic properties of naturally produced clear speech at normal speaking rates. *Journal of the Acoustical Society of America, 115,* 362–378.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America, 49,* 467–477.

Liu, S., Del Rio, E., Bradlow, A. R., & Zeng, F. G. (2004). Clear speech perception in acoustic and electric hearing. *Journal of the Acoustical Society of America, 116,* 2374–2383.

McAdams, S. (1989). Segregation of concurrent sounds: I. Effects of frequency modulation coherence. *Journal of the Acoustical Society of America, 86,* 2148–2159.

Moore, B. C. J., Peters, R. W., & Glasberg, B. R. (1993). Detection of temporal gaps in sinusoids: Effects of frequency and level. *Journal of the Acoustical Society of America, 93,* 1563–1570.

Newman, R. S. (2003). Prosodic differences in mothers' speech to toddlers in quiet and noisy environments. *Applied Psycholinguistics, 24,* 539–560.

Newman, R. S., and Hussain, I. (2006). Changes in preference for infant-directed speech in low and moderate noise by 4.5-to 13-month-olds, *Infancy,* 10, 61–76.

Newman, R. S., Weppelman, T., & Hussain, I. (2003). Does an infant-directed speaking style aid in the separation of different streams of speech? *Journal of the Acoustical Society of America, 113,* 2296.

Nozza, R. J., Rossman, R. N. F., Bond, L. C., & Miller, S. L. (1990). Infant speech-sound discrimination in noise. *Journal of the Acoustical Society of America, 87,* 339–350.

Nozza, R. J., Wagner, E. F., & Crandell, M. A. (1988). Binaural release from masking for a speech sound in infants, preschoolers, and adults. *Journal of Speech and Hearing Research, 31,* 212–218.

Olsho, L. W., Koch, E. G., Carter, E. A., Halpin, C. F., & Spetner, N. B. (1988). Pure-tone sensitivity of human infants. *Journal of the Acoustical Society of America, 84,* 1316–1324.

Patterson, R. D., & Moore, B. C. J. (1986). Auditory filters and excitation patterns as representations of frequency resolution. In B. C. J. Moore (Ed.), *Frequency selectivity in hearing* (pp. 123–178). London: Academic.

Payton, K. L., Uchanski, R. M., & Braida, L. D. (1994). Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing. *Journal of the Acoustical Society of America, 95,* 1581–1592.

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1985). Speaking clearly for the hard of hearing: 1. Intelligibility differences between clear and conversational speech. *Journal of Speech and Hearing Research, 28,* 96–103.

Picheny, M. A., Durlach, N. I., & Braida, L. D. (1986). Speaking clearly for the hard-of-hearing: 2. Acoustic characteristics of clear and conversational speech. *Journal of Speech and Hearing Research, 29,* 434–446.

Taylor, M. M., & Creelman, C. D. (1967). PEST: Efficient estimates on probability functions. *Journal of the Acoustical Society of America, 41,* 782–787.

Trehub, S. E., Schneider, B. A., & Endman, M. (1980). Developmental changes in infants' sensitivity to octave-band noises. *Journal of Experimental Child Psychology, 29,* 282–293.

Uchanski, R. M., Choi, S. S., Braida, L. D., Reed, C. M., & Durlach, N. I. (1996). Speaking clearly for the hard of hearing: 4. Further studies of the role of speaking rate. *Journal of Speech and Hearing Research, 39,* 494–509.

Werner, L. A. (1995). Observer-based approaches to human infant psychoacoustics. In G. M. Klump, R. J. Dooling, R. R. Fay, & W. C. Stebbins (Eds.), *Methods in comparative psychoacoustics* (pp. 135–146). Boston: Birkhauser Verlag.

Werner, L. A., & Bargones, J. Y. (1991). Sources of auditory masking in infants: Distraction effects. *Perception and Psychophysics, 50,* 405–412.

Werner, L. A., & Boike, K. (2001). Infants' sensitivity to broadband noise. *Journal of the Acoustical Society of America, 109,* 2101–2111.

Zeng, F. G., & Liu, S. (2006). Speech perception in individuals with auditory neuropathy. *Journal of Speech Language and Hearing Research, 49,* 367–380.

Zwicker, E., & Fastl, H. (1990). *Psychoacoustics: Facts and models.* Berlin: Springer-Verlag.