Infant Discrimination of Spectrally Weighted Speech

Kitamura, C., Gregory, J., and Kuan, S.

MARCS Auditory Laboratories,
University of Western Sydney, Australia
c.kitamura@uws.edu.au

Abstract

In this study we tested 24- and 36-week-old infants' ability to discriminate normal speech from speech with a positive or negative spectral tilt: 9dB/octave in Experiment 1, and 6dB/octave in Experiment 2. In Experiment 1 both age groups could discriminate speech with no spectral tilt from speech with positive and negative 9dB/octave emphasis. In Experiment 2 the 36-week-olds could only discriminate speech with positive 6dB/octave from normal speech, and 24-week-olds could only discriminate speech that was negatively tilted at 6dB/octave from normal speech. The diminished ability by 36-week-old infants to discriminate low frequency emphasis is attributed to a bias to high frequency information while attunement to native language consonants, and the ability to process segmental information emerges around this age.

1. Introduction

With the advent of otoacoustic emission testing, and the application of this technology to universal hearing screening programs for newborns around the world, it is possible to diagnose hearing impairment and fit hearing aids soon after birth. To date, however, hearing aids for infants have followed adult-derived amplification rules, despite evidence showing infants perceive speech differently to adults and children, and the ongoing changes to speech perception in the first year and beyond. Consequently there is a pressing need to marry the recent advances in technology and availability of newborn screening procedures, with investigations of the amplification characteristics that would optimise speech perception in young hearing-impaired infants.

Audition is the first sensory system to become functional in the unborn infant, emerging around the fifth to sixth month of foetal life [1]. The unborn child can hear aspects of both the intra-uterine and extra-uterine world [2] with the walls of the mother's body acting as a low-pass filter, only transmitting low frequency information [3]. Due to *in utero* exposure to speech, infants are born with a preference for their native language [4] and their own mother's voice [5]. Tests with low-pass filtered speech show that these early preferences are based on low-frequency suprasegmental (below 500 Hz) rather than higher frequency segmental information [4, 6].

Low-frequency prosodic information continues to be integral to infants' listening preferences. Both newborn [7] and older [8, 9] infants prefer infant-directed (ID) over adult-directed (AD) speech with this preference based on particular prosodic properties: Fernald and Kuhl [10] found infants' ID speech preferences to be based upon the fundamental frequency (pitch) of ID speech rather than its amplitude or durational characteristics. More recently, though, positive affect in ID speech has been shown to be a critical determinant of infants' ID speech preferences [11, 12]. While the suprasegmental information in ID speech appears to be important for gaining and maintaining infant attention [8, 9], communicating affect [11, 12], and facilitating social interaction [13, 14], research also shows that prosodic characteristics in ID speech and in speech generally are used by infants to extract native language information. At six months, infants prefer the dominant stress pattern of their native language [15, 16], and native over non-native language words when languages are prosodically different. In addition, exaggerated prosody facilitates their recognition of lexical [17], and syntactic [18] speech units. However, by 9 months, infants' preferences reflect the ability to integrate segmental and suprasegmental information [19, 20] and to recognize their native language via phonetic information [21].

Infants with normal hearing perceive any segmental contrast experimenters like to test them on, even those irrelevant in their language environment [22, 23]. In the months to come, infants' discrimination performance is selectively attenuated for irrelevant non-native segments, at around 6 months for vowels [24, 25], and between 7 and 11 months for consonants [26]. Therefore, to learn to attend selectively to just those contrasts in their own language, infants need exposure both to the relatively low-frequency information contained in vowels, and to the high frequency information in consonants.

Although infant speech preferences are based on low-frequency suprasegmental information, infants are good at discriminating high frequency information in speech and complex tones [27]. At present, hearing aids are fitted to hearing-impaired infants based primarily on spectral shaping found to be best for speech intelligibility for children and adults but the evidence is silent with respect to whether this is optimal for infants. Indeed it is not even known which spectral tilts infants can discriminate using speech stimuli. To date, the only research on spectral tilt discrimination comes from studies using complex tones rather than speech: Clarkson [28] found that 7-month-old infants discriminate rising and falling spectral slopes in complex tones; and Tsang and Trainor [29] found that 8-month-olds discriminate between different spectral slopes of complex tones, but only within the range of slopes found in human speech and music.

The normal speech spectrum has a slight negative spectral tilt (3 to 5 dB per octave above 800 Hz) over the amplitude spectrum. In the studies described here the long-term average spectrum of speech stimuli will be linearly shaped with a negative spectral tilt (amplitude emphasis at low frequencies), or a positive spectral tilt (amplitude emphasis at high frequencies) in a manner similar to that found in hearing aids. These manipulations will be used to investigate whether

24- and 36-week old infants can discriminate speech that has been spectrally shaped with tilt factor values of 0, –6, +6, -9, +9 dB per octave.

2. Method and Results

Two experiments were conducted examining 24- and 36-week-old infants' ability to discriminate normal speech from speech with a positive or negative spectral tilt. In Experiment 1 infants' ability to discriminate speech with a positive or negative 9dB per octave spectral tilt from normal speech was examined, and in Experiment 2 the degree of spectral tilt was reduced to 6dB per octave.

2.1 General Method

Infants were tested seated on their parent's lap facing a video monitor positioned slightly to the infant's right. An infant controlled habituation procedure was used in which the habituation stimulus was presented on repeated trials until there was an average 50% decrement in looking times over two trials compared to the average of the first two trials. Once this criterion was met, two no-change control trials followed and then infants were presented with two test trials of the novel stimulus. Infants were deemed to have discriminated spectral tilts if they showed a recovery response (longer looking times) in the novel trials compared with the control trials. To ensure that any lack of a recovery response was not the result of fatigue, infants were played a short entertaining video at the beginning and end of testing. If this did not result in recovery of visual fixation then infants were deemed to have lost interest in the procedure, not just the habituation stimulus and their data were not used in the study.

The audio and visual display for each trial began when the infant looked to an attention-getting stimulus (large flashing red, blue, pink dots) for more than 1 sec. Trials ended when the infant looked away from the visual target for more than 1.5 sec. The experimenter sat in the sound-isolated control room and monitored the infant's looking behaviour on a television monitor connected to a camera focused on the infant in the experimental room. The experimenter was unaware of the order of presentation of speech stimuli, and recorded the infant visual fixations by pressing the space bar on the keyboard with input to purpose-written software which controlled presentation of the speech stimuli, and recorded the duration of visual fixations.

The speech stimuli consisted of a female speaker's productions of multiple tokens of the utterance "swish boom" in an interesting non-monotonic speech style. "Swish boom" was chosen because the word 'swish' contains mainly high-frequency information in its consonants (/s/, /sh/) and vowel (/i/), whereas the word 'boom' contains predominantly low-frequency information in its consonants (/b/, /m/) and vowel (/u/). Levels of spectral tilt used in this study are zero (0dB**)**, positive (+6dB/octave and +9dB/octave emphasis applied to normal speech from 250 to 4000Hz) and negative (-6 dB/octave and –9dB/octave emphasis applied to normal speech from 250-4000Hz). There were six tokens in each speech stimulus set, each approximately 1.5 secs in duration, with pauses between tokens being approximately 400 ms. The presentation of stimuli in each experiment was counterbalanced, so that half the infants were habituated to speech with zero spectral tilt and tested with one of the positive or negative spectral tilts, and the other half were habituated to a positive or negative spectral tilt and tested on zero tilt.

Table 1: Participant details

	Tilt	**Infant nos**	**M age (wks)**	**Age Range**
Exp. 1 9 dB	-9dB	24	24.01	21.7-26.4
	+9dB	28	24.10	21.7-26.3
	-9dB	22	35.98	34.1-38.1
	+9dB	25	35.74	34.1-38.0
Exp. 2 6 dB	-6dB	23	24.44	22-26.6
	+6dB	24	24.33	22-26
	-6dB	25	35.99	33.6-38.6
	+6dB	24	35.84	33.6-38.4

2.2 Experiment 1: 9dB/octave spectral tilt.

In Experiment 1, two groups of 24-week-old, and two groups of 36-week-old infants were tested. At each age one group was tested with speech with no spectral tilt (0dB) and +9dB/octave spectral tilt and one group with 0dB and –9dB/octave spectral tilt. Numbers of infants and age details for each experiment are shown in Table 1. Another eight 24-week-old and three 36-week-old infants did not complete testing due to inattention or crying.

2.3 Results

The dependent variable was average looking time in each trial type (habituation, control, novel). Planned orthogonal contrasts within two (3) x 2 x 2 analyses of variance (ANOVA), one for –9 dB and one for +9 dB/octave were used to test differences between trial types as follows: (i) the first two habituation (FH) and the last 2 habituation (LH) trials, and (ii) the LH and the 2 control trials (CL). Trial type FH, LH and CL trials was a within subjects factor, and habituation stimulus

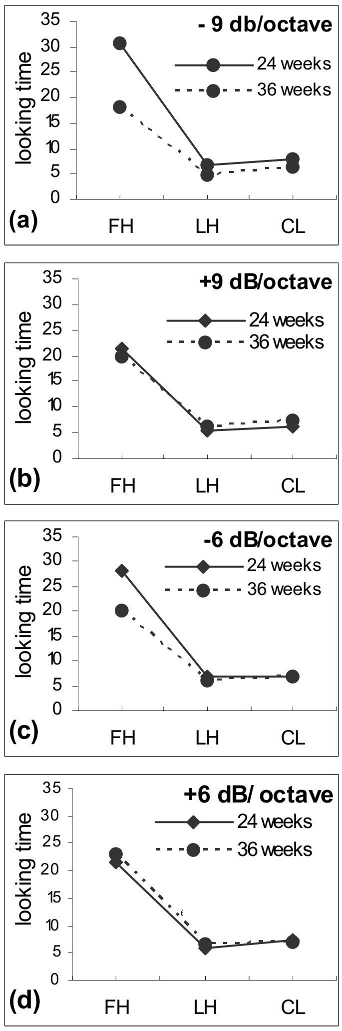

Figure 1: Means for FH, LH and CL trials, for –9dB, panel (a); +9dB, panel (b); -6dB, panel (c); and +6dB, panel (d).

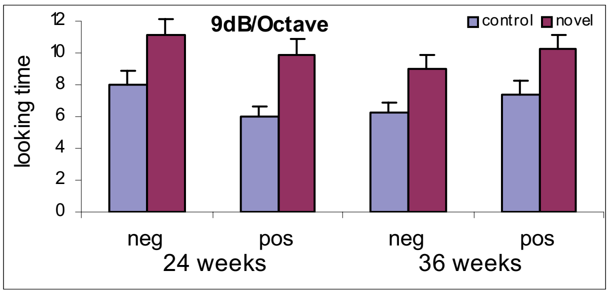

Figure 2: Mean looking times of 24- and 36-week-old infants in control and novel trials for – and +9dB/octave

(spectral tilt in habituation, normal speech in habituation) and age were between subjects factors. The results are shown in Figure 1 (a) and (b). For –9dB/octave, as is often the case in habituation procedures, there was a tendency for the younger 24-week-old infants to look longer overall than the 36-week-olds, p=.054, and the almost significant interaction between age and the FH-LH difference (p=.06) suggests that 24-week-olds looked longer than 36-week-olds at the beginning (FH) than the end (LH) of habituation. There was a significant decrease between FH and LH trials, F(1,42)=46.8, p<.0001, and a small but significant increase in looking times from LH to CL trials, F(1,42)=4.3, p<.04. For the +9dB/octave condition there was no effect of age. There was a significant decrease from FH to LH trials, F(1,49) = 70.9, p<.0001, but no other significant main effects or interactions.

To analyse recovery from habituation a (2) x 2 x 2 x 2 (trial-type x habituation stimulus x age x tilt direction) ANOVA was conducted, with trial type (control, novel) as the within subjects factor, and age, habituation stimulus (0dB or 9dB) and tilt direction (+ or -) as between subjects factors. The results are shown in Figure 2. There was a significant main effect for trial type (F (1,91)=42.19, p<.00001, η^2=.32), but no other significant main effects or interactions. Paired t-tests showed both age groups recovered to the novel stimulus irrespective of whether the habituation stimulus was 0 dB or 9 dB per octave (all p <.003). These results indicate that both 24- and 36-week-old infants can discriminate speech with no spectral tilt from speech with a positive or a negative 9 dB per octave spectral tilt.

2.4 Experiment 2: 6 dB/octave spectral tilt

In Experiment 2, 24- and 36-week-old infants were habituated to speech with no spectral tilt and tested with +6dB (or vice versa), or habituated to 0dB and tested with –6dB spectral tilt (or vice versa). The age detail and numbers of infants in each group are shown in Table 1. There were six 24-week-old and five 36-week-old infants who did not complete testing due to inattention or crying.

2.5 Results

In the 6dB experiment, two (3) x 2 x 2 ANOVAs, one for –6dB and one for +6dB/octave, with trial type (FH, LH, CL) as the within subjects factor, and habituation stimulus (0dB or 6dB) and age as the between subjects factors were conducted. The results are shown in Figure 1, panels (c) and (d). For

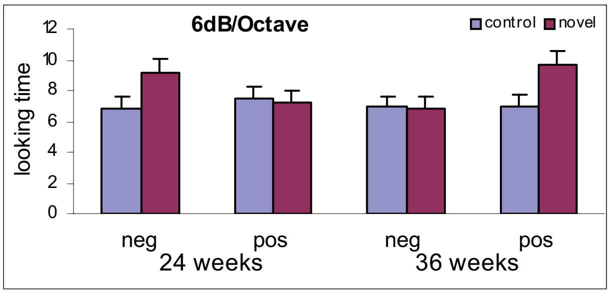

Figure 3: Mean looking times of 24- and 36-week-old infants in control and novel trial for – and + 6dB/octave

–6dB/octave, the FH trials had significantly longer looking times than the LH trials F(1,43)= 78.3, p<.0001. There were no other significant effects. The results for +6dB/octave revealed a significant decrease in looking times from FH to LH trials, F(1,44) =69.7, p<.0001; and a slight but significant increase from LH to CL trials, F(1,44)=4.7, p<.04. No other results were significant.

To analyse recovery from habituation, a (2) x 2 x 2 x 2 (trial-type x habituation stimulus x age x tilt direction) ANOVA was conducted. The results are shown in Figure 3. There was a significant main effect for trial type, F (1,87)=4.7, p<.003, η^2=.05, and a significant interaction between trial type, age and tilt direction, F (1,87)=4.2, p<.04, η^2 =.046. Follow-up paired t-tests showed 36-week-olds recovered in the *positive* 6dB/octave condition, t (23)=2.07, p<.02, but not in the negative 6dB/octave condition while 24-week-olds recovered in the *negative* 6dB/octave condition, t (22) = 2.08, (p<.02, but not in the positive 6dB/octave condition.

In summary, the results show that younger infants appear to have an attentional bias to speech with low frequency emphasis, and by 36-week-olds, this bias has diminished such that older infants do not show a recovery response in the –6dB per octave condition.

3. Discussion

Here perceptual discrimination was assessed with band-limited (.25-4 kHz) utterances, which were spectrally shaped to simulate the frequency response of hearing aids, with the tilt of the amplitude spectrum shaped to values of -6, +6, -9, +9dB per octave. Adults easily discriminate small changes in the spectral envelope (Green, 1983) but little is known about infants' developing sensitivities to spectral tilt. The results show that both 24- and 36-week old infants can discriminate no tilt from shaped amplitude envelopes of 9dB per octave, irrespective of whether tilt direction is positive or negative. However, when the degree of spectral tilt is reduced to 6dB per octave, 24-week-old infants only discriminate between speech with no spectral tilt and *negative 6dB per octave* while 36-week-olds only discriminate no spectral tilt and *positive 6dB per octave*. It seems that by the time infants reach 36 weeks, the ability to discriminate speech with low frequency emphasis has diminished, but the ability to discriminate speech with emphasis at high frequencies has improved.

Evidence shows that experience with the patterns of ambient speech systematically influences infant speech perception well before the production of words. Notably, around 24 weeks infants attune to vowels [24, 25], and here it was found that at this age they are able to discriminate small differences in low frequency spectral weighting. On the other hand around 36 weeks infants attune to consonants [26], and it was found that at this age they discriminate differences in high frequency weighting. It has also been found that between 24 and 36 weeks there is a transition from the use of predominantly prosodic suprasegmental cues to process speech to being able to integrate segmental and phonotactic information [19-21]. This emerging sensitivity to phonetic information around 9 months may account for 36-week-old infants' decreased discrimination performance for speech with low frequency emphasis. Perhaps as phonetic information begins to become important for infants, there is a relative increase in the perceptual salience (and thus discrimination performance) for high frequency information, and a corresponding decrease in the perceptual salience (and thus discrimination performance) for information with a low frequency emphasis. This may only be a brief attentional bias, which allows the infant to focus on the critical stimulation in this transition period, which is corrected once infants become more competent in using the perceptual information in the segmental aspects of speech. Further work with older infants may clarify this issue. While we can speculate on the reasons for older infants' inability to discriminate between speech with no tilt and speech with +6dB per octave spectral tilt, without doubt infants emerging sensitivity to spectral shaping of speech signals is worthy of closer examination.

4. Acknowledgements

This research was funded by a UWS Grant, and an ARC Grant (DP 0559134). We also thank Harvey Dillon from the NAL, Sydney, Australia who processed the audio files with the specified spectral tilt characteristics.

5. References

[1] G. Bredberg, The anatomy of the developing ear, in *Auditory Development in Infancy*, S. Trehub &B. Schneider, Eds. NY, London: Plenum Press, 1985.

[2] D. Querleu and K. Renard, Auditory perception in the human fetus, *Medecine et Hygiene*, 39, 2101-2110, 1981.

[3] E. M. Ockleford, M. A. Vince, C. Layton, and M. R. Reader, Response of neonates to parents' and others' voices, *Early Hum Dev*, 18, 27-36, 1988.

[4] J. Mehler, P. Jusczyk, G. Lambertz, N. Halsted, J. Bertoncini, and C. Amiel-Tison, A precursor of language acquisition in young infants, *Cogn.*, 29, 143-178, 1988.

[5] A. J. DeCasper and W. P. Fifer, Of human bonding: Newborns prefer their mothers' voice, *Science*, 208, 1174-1176, 1980.

[6] J. Mehler, J. Bertoncini, M. Barriere, and D. Jassik-Gerschenfeld, "Infant recognition of the mother's voice," *Nature*, 7, 491-497, 1978.

[7] R. Panneton Cooper and R. N. Aslin, Preference for infant-directed speech in the first month after birth, *Child Dev*, 61, 1584-1595, 1990.

[8] A. Fernald, Four-month-old infants prefer to listen to motherese, *Inf Beh and Dev*, 8, 181-195, 1985.

[9] J. F. Werker and P. J. McLeod, Infant preference for both male and female infant-directed talk: A developmental study of attentional affective responsiveness, *Can. J of Psych*, 43, 230-246, 1989.

[10] A. Fernald and P. Kuhl, Acoustic determinants of infant preference for motherese speech, *Inf. Beh. and Dev.*, 10, 279-293, 1987.

[11] C. Kitamura and D. Burnham, The infant's response to vocal affect in maternal speech, in *Advances in Infancy Research*, 12, C. Rovee-Collier, Ed., 1998, 221-236.

[12] L. Singh, J. L. Morgan, and C. T. Best, Infants' listening preferences: Baby talk or happy talk?, *Infancy*, 3, 365-394, 2002.

[13] C. E. Snow, The development of conversation between mothers and babies, *J of Child Lang*, 4, 1-22, 1977.

[14] C. E. Snow, Understanding social interaction and language acquisition: Sentences are not enough," in *Human Interaction*, M. H. Bornstein and J. S. Bruner, Eds. New Jersey: Lawrence Erlbaum Associates, 1989.

[15] P. W. Jusczyk, A. Cutler, and N. J. Redanz, Infants' preference for the predominant stress patterns of English words," *Child Dev*, 23, 648-654, 1993.

[16] L. Bosch and N. Sebastian-Galles, The role of prosody in infants' native-language discrimination abilities: The case of two phonologically close languages," 5th Eur Conf on Speech Comm and Tech, Rhodes, Greece, 1997.

[17] R. Karzon, Discrimination of polysyllabic sequences by one to four month infants, *Jof Exp Child Psych*, 39, 1985.

[18] D. G. Kemler Nelson, K. Hirsh-Pasek, P. W. Jusczyk, and K. Wright Cassidy, How the prosodic cues in motherese might assist language learning, *J of Child Lang*, 16, 1989.

[19] J. L. Morgan and J. R. Saffran, Emerging integration of sequential and suprasegmental information in preverbal speech segmentation, *Child Dev*, 66, 911-936, 1995.

[20] D. Burnham, C. Kitamura, and V. Lancuba, The development of linguistic attention in early infancy: the role of prosodic and phonetic information, in *Proc of the XIVth ICPS*, U of Cal, Berkley, 1999, 1197-1200.

[21] P. W. Jusczyk, A. D. Friederici, J. M. Wessels, V. Svenkerud, and A. M. Jusczyk, Infants' sensitivity to the sound patterns of native language words, *J. of Mem. and Lang.*, 32, 402-420, 1993.

[22] R. N. Aslin, P. W. Jusczyk, and D. B. Pisoni, Speech and auditory processing during infancy, in *Handbook of child psychology*, vol. 2. Cognition, perception and language, D. Kuhn and S. R. Siegler, Eds. Wiley: New York, 1997.

[23] D. Burnham, Developmental loss of speech perception: Exposure to and experience with a first language, *App. Psycholinguistics*, 7, 207-240, 1986.

[24] P. Kuhl, K. Williams, and A. Meltzoff, Cross-modal speech perception in adults and infants using nonspeech auditory stimulus, *J of Exp Psych*, 17, 829-840, 1991.

[25] L. Polka and J. F. Werker, Developmental changes in perception of nonnative vowel contrasts, *J of Exp Psych: Hum Perc and Perf*, 20, 421-435, 1994.

[26] J. F. Werker and R. C. Tees, Influences on infant speech processing: Toward a new synthesis, *Ann Rev of Psych*, 50, 509-535, 1999.

[27] J. M. Sinnott, D. B. Pisoni, and R. N. Aslin, A comparison of pure tone thresholds in human infants and adults, *Inf. Beh. and Dev*, 6, 3-17, 1983.

[28] M. Clarkson, Infants' intensity discrimination: Spectral profiles, *Inf. Beh. & Dev.*, 19, pp. 181-190, 1996.

[29] C. Tsang & L. Trainor, Spectral slope discrimination in infancy: Sensitivity to socially important cues, *Inf. Beh. & Dev.*, 25, 183-194, 2002.