



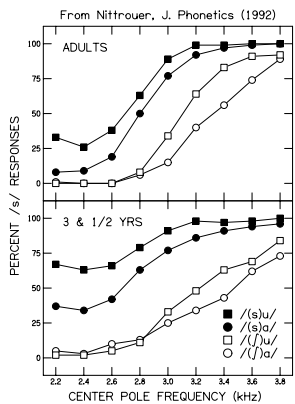
Why do children attend to global structure in speech signals?

Joanna H. Lowenstein and Susan Nittrouer
Department of Otolaryngology, Ohio State University



BACKGROUND

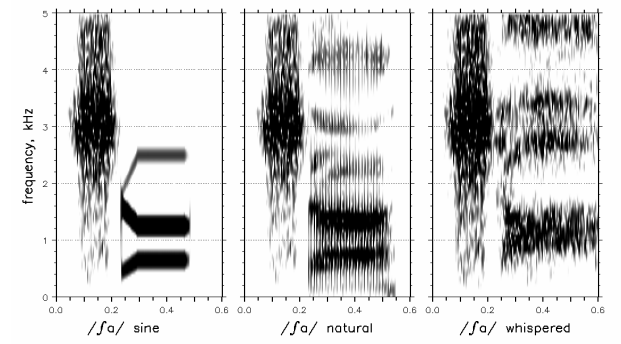
A common finding of speech perception studies is that children base linguistic decisions heavily on formant transitions. This trend is shown below: Children's labeling functions are more strongly based on vocalic formant transitions than those of adults. At the same time, children pay less attention to other signal components, such as static noises.



STIMULI

To test the hypothesis that children weight transitions greatly because they share harmonic structure, we needed stimuli that retained the transitions, but eliminated harmonic structure.

Three sets of stimuli were constructed.



The same fricative noises were used for all 3 sets: single-pole noises varying along a continuum from 2.2 to 3.8 kHz, in 200-Hz steps.

Vocalic portions were:

VOICED stimuli: Vocalic portions were taken from a male speaker saying /sa/, /su/, /ʃa/, and /ʃu/. This was our 'control' set.

SINE WAVE stimuli: Three sine waves were used to represent the first three formants of each vowel (/s)a/, /s)u/, /()a/, and /()u/).

WHISPERED stimuli: Whispered vocalic portions of /sa/, /su/, /ʃa/, and /ʃu/ taken from a male speaker.

METHOD

PARTICIPANTS

Participants were adults, 7-year-olds, and 5-year-olds, divided into two groups. One group heard the VOICED and SINE WAVE stimuli; the other group heard the VOICED and WHISPERED stimuli.

PROCEDURES

Each stimulus was presented ten times in random order. Stimuli for each vowel and stimulus type were presented separately.

Two drawings were used to represent the response label in each experiment. Each participant was required to respond by both repeating the stimuli and pointing to the appropriate drawing.

RESULTS

The figure to the right shows labeling functions for adults and 5-year-olds for the VOICED, SINE WAVE, and WHISPERED conditions.

- Mean performance for the VOICED stimuli was the same for listeners in both the SINE WAVE and WHISPERED groups.

- Adults performance across all conditions was the same.

- Children's attention to formant transitions was the same for the VOICED and SINE WAVE conditions, but was attenuated in the WHISPERED condition. Additionally, only half of the 5-year-olds were able to complete the WHISPERED condition, while nearly all of the 5-year-olds were able to complete the SINE WAVE condition.

Table 1 presents Pearson correlations for the formant transitions for adults, 7-year-olds, and 5-year-olds. Children greatly reduced their attention to formant transitions in the WHISPERED condition, while adults paid less attention to formant transitions in the SINE WAVE condition.

Table 2 presents Pearson correlations for the fricative noises for adults, 7-year-olds, and 5-year-olds. All listeners paid the same amount of attention to the fricative noise across conditions.

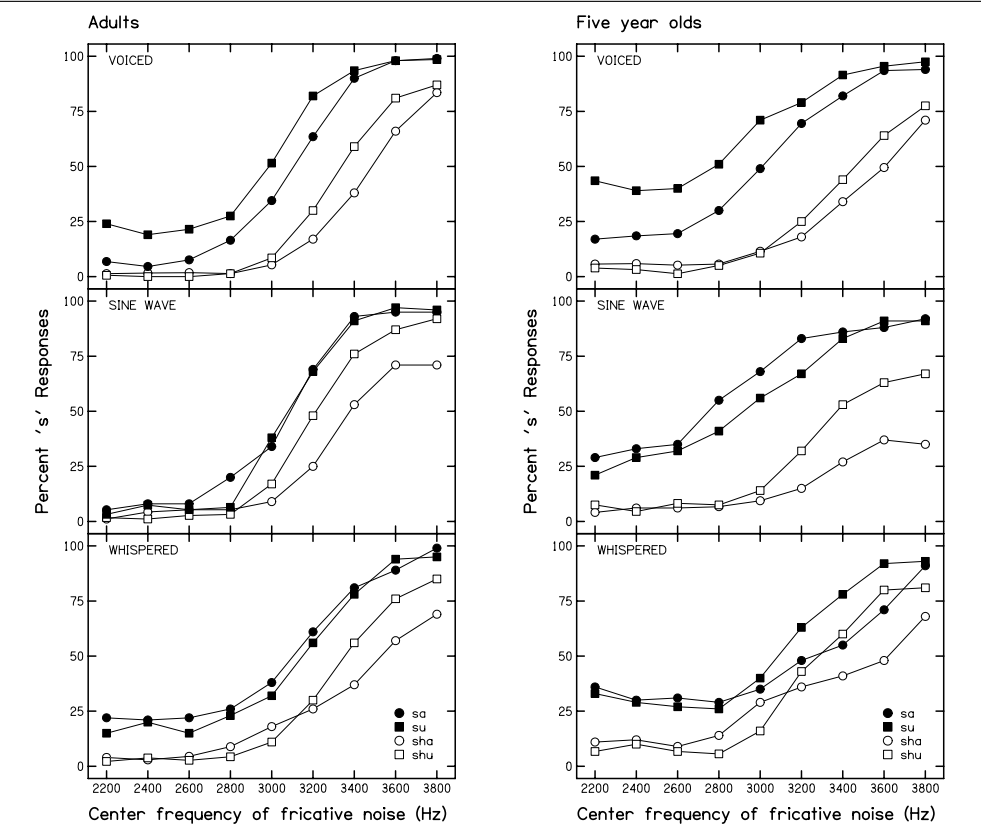


Table 1. Pearson correlations for each condition for formant transitions.

	VOICED	SINE WAVE	WHISPERED
Adults	.32 (.15)	.21 (.10)	.35 (.26)
7-year-olds	.32 (.17)	.31 (.16)	.21 (.18)
5-year-olds	.45 (.16)	.50 (.19)	.35 (.22)

Table 2. Pearson correlations for each condition for fricative noise.

	VOICED	SINE WAVE	WHISPERED
Adults	.80 (.12)	.85 (.06)	.75 (.15)
7-year-olds	.81 (.11)	.80 (.12)	.81 (.12)
5-year-olds	.68 (.14)	.65 (.20)	.69 (.18)

CONCLUSIONS

- Children attended to formant transitions as much in the SINE WAVE condition as in the VOICED condition, leading to rejection of the hypothesis that children prefer formant transitions because they are signal components that adhere to principles of ASA.

- Children did pay somewhat less attention to transitions in the WHISPERED condition, and that condition was difficult for 5-year-olds to complete. Children seemed to prefer the SINE WAVE stimuli, perhaps because the stimuli provided a greater degree of global spectral structure than the WHISPERED stimuli did. This is commensurate with our finding that children can process SINE WAVE sentences sooner than they can AMPLITUDE ENVELOPE sentences. (see Poster 22)

REFERENCES

Bregman, A. S. (1990) *Auditory Scene Analysis*. Cambridge, MA: MIT Press.

MacNeilage, P. F. and Davis, B. (1991). Acquisition of speech production: Frames, then content. In M. Jeanerod (Ed.), *Attention and Performance XIII* (pp. 453-476). New York: Erlbaum.

Menn, L. (1978). Phonological units in beginning speech. In A. Bell and J. B. Hooper (Eds.), *Syllables and segments* (pp. 157-172). Amsterdam: North-Holland Publishing Company.

Menn, L. (1983). Development of articulatory, phonetic and phonological capabilities. In B. Butterworth (Ed.), *Language production: Development, writing and other language processes* (pp. 3-50). New York: Academic Press.

Nittrouer, S. (1992). Age-related differences in perceptual effects of formant transitions within syllables and across syllable boundaries. *J. Phonetics* 20, 351-382.

Vihman, M. M. and Velleman, S. L. (1989). Phonological reorganization: A case study. *Language and Speech* 32, 149-170.

ACKNOWLEDGEMENT

Work supported by NIDCD Grant No. DC-00633

Such findings have been interpreted as evidence that children initially focus on "global" structure; that is, structure that is supra-phonetic. This notion finds support in studies of speech production, where it is widely held that children reveal mastery first of supra-phonetic gestures. Terms coined to convey this concept include **word recipes** (Vihman & Velleman, 1989), **articulatory routines** (Menn, 1978; 1983), and the **frame-content** model (MacNeilage & Davis, 1991).

But vocalic formant transitions happen to share qualities that are described by Auditory Scene Analysis (ASA; Bregman, 1990). In particular, they share the same **harmonic structure**. So, the question arises:

Do children attend to signal attributes that conform to ASA?