visiBabble for Reinforcement of Early Vocalization

Harriet Fell
College of Computer and Information Science
Northeastern University
Boston, MA, 02115, USA
+1 617 373 2198
fell@ccs.neu.edu

Cynthia Cress
Communication Disorders Department,
University of Nebraska
Lincoln, NE, 68583-0732, USA
+1 402 472 4431
ccress1@unl.edu

Joel MacAuslan
Speech Technology and Applied Research
54 Middlesex Turnpike
Bedford, MA, 01730, USA
+1 781 276-4540
JoelM@S-T-A-R-corp.com

Linda Ferrier
Speech-Language Pathology and Audiology
Northeastern University
Boston, MA, 02115, USA
+1 617 373 2892
l.ferrier@neu.edu

Gwen Sterup
Communication Disorders Department,
University of Nebraska
Lincoln, NE, 68583-0732, USA
+1 402 472 4431
cub725@yahoo.com

Andrea Heinrich
Communication Disorders Department,
University of Nebraska
Lincoln, NE, 68583-0732, USA
+1 402 472 4431
andreajheinrich@yahoo.com

ABSTRACT

The visiBabble system responds with animations to an infant's syllable-like productions and records the acoustic-phonetic analysis. The system reinforces production of syllabic utterances associated with later language and cognitive development. We report here on the development of the visiBabble prototype and field-testing with children at risk for being nonspeaking.

BACKGROUND

Physical, neurological, oral/motor, and cognitive impairments can all significantly impair speech. A child may not be able to produce a sound when he or she wants to, may produce a limited range of sounds, or may not have learned to associate his or her sounds with meaningful referents [2]. Infant vocalizations are effective predictors of later articulation and language abilities. Research studies emphasize the importance of early speech intervention for children at risk for being non-speaking. They also point out the difficulty of providing sufficient speech practice and feedback for children with such atypical speech patterns through traditional forms of intervention and interaction.

The visiBabble system responds to a child's syllable-like productions with brightly colored animations and records the acoustic-phonetic analysis. The system reinforces the production of syllabic utterances that are associated with later language and cognitive development. As a child interacts with visiBabble, the program collects and analyzes the infant's utterances so that it can be used as either a toy or clinical tool. The techniques involved with VisiBabble directly reflect best practice in prespeech intervention demonstrated in treatment efficacy studies. VisiBabble uses a responsive strategy based on a commonly accepted "play" mode of speech intervention, in which activities are child-initiated, child-focused, and do not include specific training events [13].

Central to the acoustic-phonetic processing in visiBabble are *landmarks*, points in an utterance around which listeners extract information about the underlying distinctive features. They mark perceptual foci and articulatory targets [13]. As babbling develops, the infant begins to coordinate control of the vocal folds and the velopharyngeal opening with control of the tongue blade and the lips, and the true consonants appear. In the landmark model, the larynx and the velum are considered secondary articulators, and they are "bound" to control by the primary articulators. That is, implementation of the laryngeal and nasal features depends, in some ways, on the implementation of the primary articulator.

METHODOLOGY

The visiBabble prototype system includes a modern notebook computer, a microphone, a 15" flat-panel display, and software. The visiBabble prototype responds to the child's utterances with five different brightly colored animations that cycle to avoid habituation: (a train, a bird, a frog and two cartoon creatures that move across the screen.. It responds to the start of each syllable it detects by advancing the current animation one step.

It determines that a syllable has started either by noting the onset of voicing or by a voiced closant that occurs at least 100 ms after the start of the previous syllable. Admittedly, a syllable might start with a burst before the voicing onset but, to avoid responding to noise, visiBabble waits for the onset of voicing. The system responds in no more than 0.25 second of the corresponding acoustic event.

As visiBabble runs, it makes a digital recording of the session which it saves in wav format. It also saves a record of the times and types of landmarks it found during the session. Data is collected during all phases of all formats to allow a comparison of behavior during the baseline and active phases.

After a visiBabble session, a post-processor uses the landmark types and times recorded by visiBabble to tabulate the average number of syllables per utterance, as well as the number of utterances comprised of 1, 2, 3, and more syllables, for each subject and session. Greater variety of landmarks indicates greater vocal complexity.

RESULTS

We describe here the results from twenty-one sessions, with three children with severe expressive impairments and one premature child. Three completed sessions, one with each of three children, were not used in this study, due to mechanical or adult errors. In the tests reported here, the system rarely responded to noise and whispering that can be heard in the background. One additional developmentally delayed child, K5, was clearly not yet a candidate for visiBabble and the details of sessions with this child are not included in this analysis. This child vocalized fewer than ten times in six complete sessions.

Each session consisted of a 10-minute visiBabble run in A-B-A format. The no-feedback, A phases (A1 and A2) each ran for 2.5 minutes. The B phase, with feedback, ran for 5 minutes. Data was collected throughout the 10-minute session and we compared the results during the B phase with the combined results of the A1 and A2 phases. Our results, therefore, are a comparison of vocalizations during five minutes with feedback and five minutes without.

In the twenty-one sessions of children with severe expressive impairments, amount of syllable and utterance production was significantly greater in the visiBabble test (B) segments compared to the control sessions (sum of A1 and

A2). The average vocal production varied across test and control conditions: mean number of syllables: 676 test vs. 442 control, and mean number of utterances 522 test vs. 355 control. The quantity of vocalizing was significantly different at $p < .01$ between test and control conditions for number of syllables and number of utterances, using nonparametric Wilcoxon signed-rank comparisons to account for the non-normal population distribution.

This difference was demonstrated not only in the total sessions, but also for eighteen of the twenty-one sessions individually. In fact, the total syllable production rose significantly in the test (B) condition for every child, by amounts ranging from 26% to 203%. The number of utterances rose similarly.

More importantly, the number of syllable *types*, a measure of the variety of syllables produced, also rose substantially and significantly, from a mean of 9.0 in the control (A) condition to 11.9 in the test (B).

There were no statistically significant differences in the mean number of syllables per utterance (1.30 in A vs. 1.25 in B: most utterances were monosyllabic for these infants), mean syllable duration (247 vs. 241 ms), or mean number of landmarks per syllable (2.81 vs. 2.89).

These results are very encouraging. We observed an approximately 50% increase in syllable and utterance production and a 33% increase in the variety of syllable types when visiBabble was active.

ACKNOWLEDGMENTS
This work was sponsored in part by NIH Grant R41 DC005534.

REFERENCES

1. Cress, C.J., and Ball, L. Strategies for promoting vocal development in young children relying on AAC: Three case illustrations, *Proceedings of RESNA '98* (Minneapolis, MN, June 1998). RESNA Press, 44-46.
2. Fell, H.J., MacAuslan, J., Ferrier, L.J., and Chenausky, K., Automatic babble recognition for early detection of speech related disorders, *Proceedings of Assets '98* (Marina del Rey, CA, April 1998), ACM Press, 59-66.
3. Fell, H.J., MacAuslan, J., Ferrier, L.J., Worst, S.G., and Chenausky, K., Vocalization age as a clinical tool, *Proceedings of ICSLP '02,* (Denver CO, September 2002). 2345-2348.
4. Locke, J.L., Babbling and early speech: Continuity and individual differences, *First Language*, 9 (1989), 191-206.
5. McReynolds, L.V. and Kearns, K.P.. *Single Subject Experimental Designs in Communication Disorders*, University Park Press, Baltimore: MD, 1983.
6. Shriberg, L. and Kwiatkowski, J. Phonological disorders II: A conceptual framework for management, *Journal of Speech and Hearing Disorders*. 47 (1982), 242-256.
7. Stevens, K.N., Manuel, S., Shattuck-Hufnegel, S., and Liu, S., Implementation of a model for lexical access based on features, *Proceedings of ICSLP '92*, (Banff, Alberta, 1992), 1, 499-502

Student Bios

Gwen Sterup is a graduate student in speech-language pathology finishing her degree at the University of Nebraska-Lincoln in August 2004. She collected for 3 of the children in the visiBabble project and interpreted the data collected as part of her directed research project for her master's degree.

Andrea Heinrich is a graduate student in speech-language pathology finishing her degree at the University of Nebraska-Lincoln in August 2004. She collected for 3 of the children in the visiBabble project and interpreted the data collected as part of her directed research project for her master's degree.