Chapter 8

CSS and Children: Research Results and Future Directions

Kathryn D. R. Drager
The Pennsylvania State University, USA

Joe Reichle
University of Minnesota, USA

ABSTRACT

Currently, many computer-based augmentative and alternative communication (AAC) systems use speech output, either synthesized or digitized speech. The goal of this chapter is to provide a review of the research to date on computerized synthesized speech (CSS) with children. Information on the intelligibility and comprehension of CSS for children is presented, and the variables that may affect these, including context, speech rate, age of the child, the language(s) spoken by the listener, experience with CSS, and background noise. Each of these factors and the research support with child participants are discussed. The intelligibility of digitized speech is also discussed. Additionally, this chapter will address the attitudes and preferences of children regarding CSS, as well as hypotheses about the role that CSS may play for children with significant communication disabilities that require AAC. Finally, future research priorities are presented.

INTRODUCTION

Children encounter computerized synthesized speech (CSS) in a variety of places. Synthesized and digitized speech is found in educational software and computer games, as well as in augmentative and alternative communication (AAC) systems. Computerized speech in each of these applications should be as intelligible as possible to maximize educational opportunities and, for a child with significant communication disabilities, the potential for peer interaction.

Approximately 8-12 individuals per 1,000 experience speech and language impairments severe enough to significantly limit effective communication with others (Beukelman & Ansel, 1995). Many of these individuals use gestural- or graphic-based AAC systems. Approximately 12% of children receiving special education services require AAC (Binger & Light, 2006). For many of these children,
computerized AAC systems with speech output are available to support their communication, using digitized, synthesized, or a combination of digitized and synthesized speech.

Digitized speech is recorded human voice stored as sampled segments of sound waves (Schlosser, 2003). Synthesized speech is computer generated according to a set of rules in an algorithm. Different synthesizers have used different types of coding to produce speech; some are based on the human voice. There is a wealth of research on the intelligibility and listener comprehension of synthesized speech for adults under ideal listening conditions. However we know relatively little about the usefulness of synthesized or digitized speech when young children serve as listeners. For the purposes of this chapter, the term computerized synthesized speech (CSS) will encompass both synthesized and digitized speech.

Intelligible speech output has several advantages for children who require AAC. First, intelligible CSS may allow children who require AAC an opportunity to experience more naturalized interaction with peers. Speech output may be a more comfortable method of communicative exchanges for peers, who are accustomed to communicating with one another via speech. CSS may assist in providing children who require AAC with opportunities to develop critical social skills and relationships that otherwise would not be available. Second, there is some evidence that CSS may enhance learning of AAC symbols (Schlosser, Belfiore, Nigam, Blischak, & Hetzroni, 1995). Third, CSS may increase comprehension of spoken language for children learning to use AAC using naturalistic language instruction (Sevcik & Romski, 2002). Fourth, the use of CSS allows for increased independence in communication with a wide range of communication partners, such as other children who are not literate or who are visually impaired, as well as in groups or over the telephone. In addition to these reasons, there are numerous applications for speech output in educational software and other computer games. To fully realize these advantages, it is necessary to examine the intelligibility and listener comprehension of CSS and the factors that influence these outcomes. It is also necessary to determine the effects of using CSS on attitudes of listeners, as well as the preferences of children regarding these voices. Lastly, it is also important to consider the role that speech output plays for children who rely on CSS for communication.

**BACKGROUND**

A processing constraint model has driven the majority of research examining the intelligibility and listener comprehension of CSS with adults. Theoretically, humans have a finite capacity for attention (Moray, 1967), with the brain allocating these resources across tasks. Tasks that require a large amount of processing resources will be completed at the expense of other tasks. Natural speech is characterized by redundant acoustic cues (Klatt, 1987). The information from the natural speech signal is rich, and little attention needs to be allocated by the listener for speech perception. In contrast, synthesized speech contains very few of these redundant cues, requiring increased processing resources for perception (Duffy & Pisoni, 1992). Fewer resources remain for comprehension and other higher order processing, and few remain for any other demands in communication interactions. Children, however, are working within the constraints of a more limited working memory capacity than adults (Case, 1985; Dempster, 1981). The limited capacity will impact the attentional resources available for deciphering synthesized speech. Thus, it may not be possible to generalize the results of research on speech output with adults as listeners.
Related Content

Lexical Challenges in the Intersection of Applied Linguistics and ANLP
www.igi-global.com/chapter/lexical-challenges-intersection-applied-linguistics/64580?camid=4v1a

Analysis-by-Synthesis Echo Watermarking
Wen-Chih Wu and Oscal Chen (2008). Digital Audio Watermarking Techniques and Technologies: Applications and Benchmarks (pp. 152-171).
www.igi-global.com/chapter/analysis-synthesis-echo-watermarking/8330?camid=4v1a

Building Collaborations between University Pre-Service Student-Teachers and English Language Students through a Socially Mediated Network
www.igi-global.com/chapter/building-collaborations-between-university-pre-service-student-teachers-and-english-language-students-through-a-socially-mediated-network/108776?camid=4v1a

Game On! Teaching Foreign Language Online
www.igi-global.com/chapter/game-on-teaching-foreign-language-online/108770?camid=4v1a