Chapter 4
Humanizing Vox Artificialis: The Role of Speech Synthesis in Augmentative and Alternative Communication

D. Jeffery Higginbotham
University at Buffalo, USA

ABSTRACT
In this chapter, the authors will discuss the use of speech synthesis as a human communication tool in what is now referred to as Augmentative and Alternative Communication (AAC). The authors will describe the history and use of speech synthesis in AAC, relevant stakeholders, a framework for evaluating speech synthesis in AAC and relevant research and development with respect to intelligibility, comprehension, social interaction, emotional expressivity and personal identity potential of current implementations of speech synthesis into SGD technologies. Throughout the chapter recommendations will be made for making SGDs more effective and appropriate for social interaction and emotional expression. This chapter will also provide 1st person accounts of relating to SGD use in order to provide a stakeholder perspective.

INTRODUCTION
What has happened to the human voice. Vox Humana. Hollering, shouting, quiet talking, buzz. I was leaving the airport, it’s in Atlanta. You know you leave the gate, you take a train that took you to concourse of your choice, and I get in to this train. Dead silence, Few people are seated or standing. Up above you hear a voice that once was human voice, but no longer, now it talks like a machine - “Concourse 1, Fort Worth, Dallas, Lubbock” - that kind of voice. And just then the doors are about to close - pneumatic doors - one young couple rush in and push open the doors and get in. Without missing a beat, that voice above says, “Because of late entry, we’re delayed 30 seconds”. The people looked at that couple as if that couple just committed mass murder, you know. And the couple are shrinking like this, you know... I’m known for my talking. I’m gabby. And so I say, “George Orwell, your time has come and gone!” I expect a laugh - dead silence. And now they look at me and I’m
with the couple, the three of us are at the Hill of Calvary on Good Friday. And then I say, “My God, where’s the human voice”? And just then there’s a little baby - maybe the baby is about a year or something. And I say, “Sir or Madam” to the baby. “What is your opinion of the human species”? Well what does the baby do? baby starts giggling. I said, “Thank God! The sound of a human voice.” [Terkel (2008)].

Studs Terkel’s soliloquy on the human condition sets the stage for this chapter on the use of synthesized speech in computerized Speech Generating Devices (SGD) by individuals with Complex Communication Needs (CCN). For more than any other application of this technology, speech synthesis (and supporting computer tools) is charged with the responsibility to serve as a major expressive modality during social interactions. As argued in this chapter, this responsibility goes beyond that of merely being a tool to convey information, it also serves importantly as an interactive tool for achieving common ground and as a means for conveying a speaker’s, health, attitude, affiliations, emotion, meaning and identity. What can be done to humanize speech synthesis for individuals who use SGDs as their social voice? What is it about the vox artificialis that keeps it from being one’s voice?

In the course of this chapter, we will discuss the use of speech synthesis as a human communication system in what is now referred to as Augmentative and Alternative Communication (AAC). We will describe 1) how speech synthesis is currently used in Speech Generating Devices (SGDs); 2) the stakeholders in AAC; 3) a history of speech synthesis in AAC, 4) frameworks for evaluating the Augmented Voice. We will then cover AAC research and development in areas of speech intelligibility, sentence and discourse comprehension, social interaction, and emotion and identity. Throughout the chapter, recommendations are presented for improving speech synthesis devices to make them more effective and appropriate for interaction and expression. That is to say, we will be responding to Terkel’s plea to humanize vox artificialis.

**SYNTHETIC SPEECH AND AAC TECHNOLOGIES: SPEECH GENERATING DEVICES**

Over the past three decades, major technological advances in AAC area have resulted in special purpose computerized communication aids and improved microcomputer access for those individuals with significant communication and physical access challenges. These advances include customized input and computer interfaces, sophisticated vocabulary databases and search algorithms, synthetic speech output, as well as new ways to access to standard computers and the internet.

Two forms of speech output used in past and current AAC are text-to-speech synthesis and digitized speech. We will focus here on those SGDs that utilize some form of text-to-speech synthesis. Speech synthesis is inextricably linked with the SGD platform in which it is employed. The SGD features influence the output method used to speak a message (e.g., speak after each sentence, speak after each word, speak after each keypress), the prosodic capabilities and flexibility associated with message production (e.g., utterance intonation, word emphasis) and the speech modifications available to the user during face-to-face interaction (i.e. ability to change volume, speed, voice settings in real-time).

The form-factor for the modern SGD varies from handheld PDA’s, laptop PCs, and specialized platforms dedicated to AAC use (Figure 1). Users of AAC technologies produce speech output by selecting a pre-stored message or by constructing a message through individual letters and stored words and phrases. Message construction can be facilitated through word prediction,