# Comprehension of Synthetic Speech and Digitized Natural Speech by Adults with Aphasia

## Articles at a Glance · Pathways for Aphasia

Hux, K., Knollman-Porter, K., Brown, J., & Wallace, S. (2017). Comprehension of synthetic speech and digitized natural speech by adults with aphasia, Journal of Communication Disorders, 69, 15-26, DOI: 10.1016/j.jcomdis.2017.06.006

**Background Information**

People with aphasia often demonstrate long-term reliance on others to decipher and interpret written materials, thus causing them to feel frustrated and dependent. A possible solution to this problematic scenario is using assistive technology applications with text-to-speech (TTS) conversion capabilities and computer-generated (i.e., synthetic) speech output to provide simultaneous bimodal (i.e., written and auditory) presentation of content.

- Synthetic speech generation involves using a device to (a) capture text appearing in digital or analog format; (b) convert it to a digital representation, as needed; (c) transform it into corresponding phonemes and allophones; and (d) convert it from a digital signal to analog speech waveforms.
- It typically is modifiable with regard to the rate and volume of speech output.
- However, using a TTS system to generate synthetic speech production of a written text will only be effective as a compensatory reading strategy if two conditions are met:

  (a) the generated speech output is comprehensible to a person with aphasia
  (b) the individual has sufficient auditory—or combined auditory and reading—comprehension skills to understand the presented message.

The primary purpose of the current study was to compare the auditory comprehension of people with aphasia when listening to sentences generated in three ways:

(a) with digitized natural speech produced by a male speaker of American English,
(b) with the synthetic “Alex” speech available via Macintosh platform computers, and
(c) with the synthetic “David” speech available via Windows platform computers.

The secondary purpose of this research was to solicit opinions from people with aphasia about their preferences and perceptions regarding the quality of the digitized natural speech and the selected synthetic speech options.

- The focus of the study was to research the assumption that people with a variety of aphasia types and severities can comprehend current renditions of synthetic speech with accuracy levels comparable to their comprehension of natural speech presented in a digitized format.
- Researchers have documented that the foremost reason people with disabilities abandon assistive technology devices relates to the failure of practitioners to consider a user's desires and preferences ( Scherer & Glueckauf, 2005 ).

With regard to TTS system usage by people with aphasia, this means taking into consideration opinions about the naturalness, ease of understanding, and clarity of the generated speech.

**Key Findings**

- Ranking and Likert-scale rating data revealed a preference for digitized natural speech and "David" synthetic speech over "'Alex" synthetic speech. Results suggest many individuals with aphasia can comprehend synthetic speech options available on popular operating systems.
- Participants obtained the highest overall comprehension accuracy scores for the digitized natural speech condition, followed by the Alex synthetic speech, and then David synthetic speech.
- 11 of the 20 participants selected the digitized natural speech as their most preferred option after listening a second time to example sentences produced with digitized natural speech.
    - Participants' rationales for preferring digitized natural speech reflected their appreciation of the quality, naturalness, and comprehensibility of generated utterances (e.g., P1: Smoother…I like this so much better. P2: It was clear. P6: Normal. P18: I know what's going on. P20: Voice had an even tone.).
    - Rationales provided by participants preferring either the Alex or David synthetic speech option were primarily related to ease of understanding (e.g., P5: [David was] easy to understand. P11: Man's voice [Alex] was better, nicer, different. P15: [David was] easier on my ears.).
    - Participants justified their dislike of the Alex synthetic speech by making comments such as, Garbage (P6), It's harder to understand (P15), I don't like the wa wa wa (P1), Too fast (P18), and Voice was too high (P20).

Summary

- Participants considered the digitized natural speech and David synthetic speech to be comparable regarding naturalness, clarity, and ease of understanding even though they performed better on the comprehension task when listening to digitized natural speech.

- Another potential objective of future investigations is determining whether and how much people with aphasia improve their comprehension of synthetic speech with repeated exposure.
- Researchers have reported on the boost to comprehension for people with and without intellectual disabilities can be accredited to given repeated opportunities to listen to synthetic speech.
- Individuals' unique opinions and needs are important considerations affecting long-term acceptance and use of supportive technology devices (Scherer, 2005). As such, examining preferences and needs on an individual basis and balancing the relative advantages and disadvantages of currently available renditions of synthetic speech is a necessary step when pursuing the use of TTS systems as potential reading supports for people with aphasia.

**Applying the Findings with Tobii Dynavox Aphasia Pages**

- Acapela Voices: Tobii Dynavox software and devices come with natural sounding Acapela voices.

- Recordings: Easily record speech onto any of the buttons or scripts. This may help improve comprehension.