Audio Dilation in Real Time Speech Communication

John S. Novak, III, Jason Archer, Robert Kenyon
Electronic Visualization Laboratory, University of Illinois at Chicago

Valeriy Shafiro
Rush University Medical Center

Definition and Motivation

Audio Dilation is real time, distortion-free slowing of speech or any other audio signals. (1) Similar rate manipulations have been shown to improve speech perception in challenging, multi-talker environments for older adults. (2) It is hypothesized that this technique can improve perception in a wide range of applications (e.g., native/non-native speech communications; complex, multi-talker environments such as pilot/air traffic controllers, or first responder/dispatcher communications) by managing or re-allocating competition between cognitive resources.

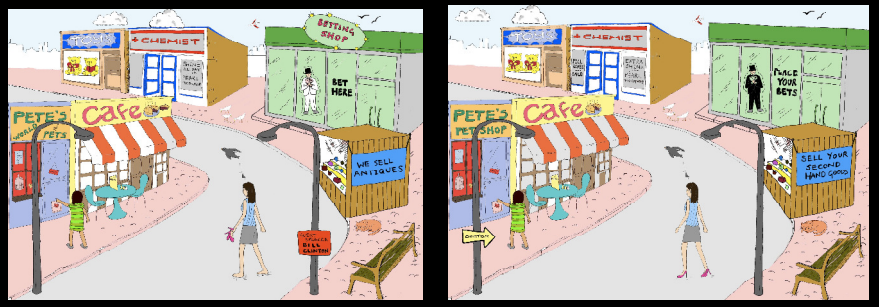
Method

Subjects: 10 young adults participated (average age = 28 (18 - 57); 6 male, 4 female).

Tasks/Procedure: Seated in different acoustically separated rooms, a pair of subjects were asked to find twelve differences between two similar Diapix (3) images as quickly as possible. Subjects conversed through headphones connected to a pair of laptop computers capable of dilating audio as generated. After completing the experiment subject pairs were debriefed and asked to complete a questionnaire to gauge their experience of dilated interactions.

Conditions: Each subject pair was tested in four conditions, each with a separate pair of Diapix images: both undilated (UU), both dilated (DD), and only one dilated (UD/DU). (First character refers to subject of analysis, e.g., "UD" means subject's speech not dilated, but subject hears dilated partner.) All conditions were randomized to attenuate possible learning effects. Dilated speech was lengthened by 40%.

Diapix Image Test



Analysis and Results

Speech produced by subject pairs was recorded in all conditions. For each condition, speech rate measurements were performed using a PRAAT script for syllable nuclei detection (4) to determine both:

- Articulation rate (syllables per second during phonation, excluding pauses)
- Number of pauses greater than 0.3 sec

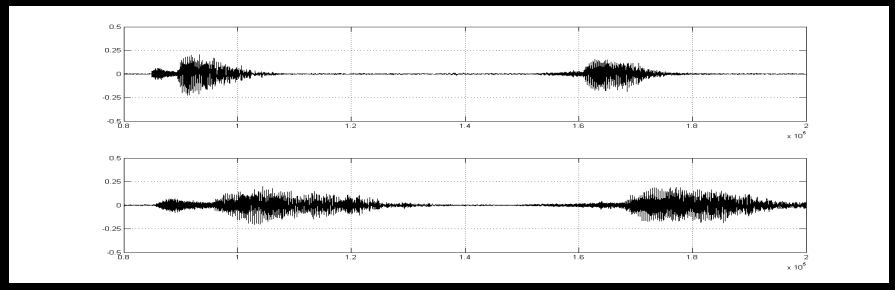
	UU	DD	UD	DU
Articulation rate:	3.94	3.96	3.98	3.99
Number of Pauses:	122.56	145.60	109.9	115.3
Total Time (in sec.):	455.69	605.08	484.57	487.79

Discussion and Conclusions

The data demonstrates:

- Audio dilation of speech can be used during communication to solve cooperative tasks,
- There is no difference in problems solving task results with or without dilation,
- Participants' rate of speech and number of pauses do not change depending on whether or not they hear regular or dilated speech addressed to them. ($p > 0.05$.)

Several participants expressed frustration with the slow rate of speech given the task's time constraints. It appears that using Audio Dilation under time pressure, participants could communicate optimally at their regular speech rate. Slowing speech had no effect but could potentially become a hindrance. However, if a task involved following a set of complicated instructions spoken quickly and/or with an accent, or was heard by a non-native, older, or hearing-impaired listener, Audio Dilation may have had a more pronounced effect.



1. Novak III, John S., et al. "Networked on-line audio dilation." Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication. ACM, 2014.
2. Gygi, Brian, and Valeriy Shafiro. "Spatial and temporal modifications of multitalker speech can improve speech perception in older adults." Hearing research 310 (2014): 76-86.
3. Van Engen, Kristin J., et al. "The Wildcat Corpus of native-and foreign-accented English: Communicative efficiency across conversational dyads with varying language alignment profiles." Language and speech 53.4 (2010): 510-540.
4. De Jong, N.H. & Wempe, T. (2009). Praat script to detect syllable nuclei and measure speech rate automatically. Behavior research methods, 41 (2), 385 – 390