
This version is available at https://strathprints.strath.ac.uk/55176/

Strathprints is designed to allow users to access the research output of the University of Strathclyde. Unless otherwise explicitly stated on the manuscript, Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Please check the manuscript for details of any other licences that may have been applied. You may not engage in further distribution of the material for any profitmaking activities or any commercial gain. You may freely distribute both the url (https://strathprints.strath.ac.uk/) and the content of this paper for research or private study, educational, or not-for-profit purposes without prior permission or charge.

Any correspondence concerning this service should be sent to the Strathprints administrator: strathprints@strath.ac.uk
Analyzing speech movement variability from audio recordings

Frits van Brenk & Anja Lowit
Division of Speech and Language Therapy, University of Strathclyde, Glasgow, UK
Contact: frits.brenk@strath.ac.uk  a.lowit@strath.ac.uk

Introduction

Measuring variability in speech motor movements
- Direct: kinematics movements by EMA, EMG, MRI and ultrasound → invasive and/or expensive
- Indirect: audio data → non-invasive and cheap

Speech parameters:
- Intensity and F0: laryngeal activity
- F1 and F2: tongue position

Variability measures:
- Spatiotemporal Index (STI), linear [1]
- Functional Data Analysis (FDA), non-linear [2]

Research question in this study:
1. What is the effect of changing speaking rate, sentence length and complexity, and performing dual tasks on variability in speech motor control?
2. Are variability measures of acoustic data comparable with variability results obtained from kinematic data in earlier studies?

Methodology

Participants
- Seventeen native Scottish speakers, 13 females and 4 males, age range 18 to 45 years (mean = 27.2 years, SD = 8.6 years).

Experimental task
- Repeat the phrase “Tony knew you were lying in bed” around 20 times.

Speaking conditions:
- Habitual speech rate (baseline condition)
- Fast rate
- Habitual rate, increased sentence length: “One two three Tony knew you were lying in bed five six seven”
- Habitual rate, increased sentence length and complexity: “I heard that Tony knew you were lying in bed this Sunday morning”
- Habitual rate with simultaneous spiral drawing

Speakers
- Audio data collected with portable voice recorder and head-mounted microphone.
- Annotation and extraction of amplitude envelope, F0, F1, and F2 tracks in Speech Filing System.
- Variability analysis of sentence repetitions with custom Matlab software [3].
- STI: linear stretching of tracks, cumulative summing of standard deviations across tracks
- FDA: non-linear stretching, spatial and temporal variability separately

Spatiotemporal Index

Parameter

Fast Habitual Slow

Increased sentence

Analysis

Accuracy of formant estimation in SFS is sometimes problematic.
- Solution: remove outliers by iterative re-assignment of peaks based on mean trajectories.

Discussion

Results

Spatial Variability

Temporal variability

Speech Rate: STIs for F0 and F1 were significantly higher in slow compared to habitual speech rate.
- Sentence complexity: STIs for F0 were significantly higher in slow speech rate.
- Increased length and complexity conditions compared to habitual speech rate.

Concurrent motor task: STIs in the dual motor task did not significantly differ from habitual speech rate.

Speech Rate: Spatial variability for Amplitude and F0 was significantly higher in slow speech rate compared to habitual and fast speech rate. For F1, spatial variability was higher at slow speech rate compared to fast speech rate.
- Sentence complexity: Spatial variability of Amplitude was significantly higher in sentences with increased length and complexity compared to the baseline sentence. For F0, variability in sentences with increased length and complexity was higher compared to the baseline sentence.

Concurrent motor task: Spatial variability of F0 was higher in the dual motor task compared to the baseline condition.

Conclusions

- Measuring variability in audio data is a promising and easy applicable method to analyse speech motor control.
- Results are similar to direct measures of variability.
- However, possible problems with data collection and processing may lead to a decrease in sensitivity compared to direct kinematic measures.

References