This is the published version of a paper presented at Nineteenth Biennial Conference on Motor Speech: Motor Speech Disorders & Speech Motor Control, Feb 22 – 25, 2018, Savannah, Georgia.

Citation for the original published paper:

How well does DDK task performance predict fluent speech articulation?
In:

N.B. When citing this work, cite the original published paper.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-144767
HOW WELL DOES DDK TASK PERFORMANCE PREDICT FLUENT SPEECH ARTICULATION?

Fredrik Karlsson¹, Marcus Karlsson¹ & Lena Hartelius²
¹Umeå University, ²Gothenburg University

ABSTRACT

Rapid repetition of CV syllables (oral diadochokinesis, DDK) puts pressure on the speech motor system in order to afford detection of a deteriorating articulatory ability.

This study assessed:
1) how well objective quantifications of DDK sequences can be used to identify PD speakers among age matched controls (NC), and
2) how well a reduced articulatory ability in PD speakers may be predicted using quantifications of DDK sequences performed in the same recording session.

The comprehensive DDK sequence quantification achieved a 93% classification accuracy for speaker group membership. A 93% accuracy for reduced fluent speech articulation quality.

METHOD

- Alternating Motion Rate (AMR) produced by 38 PD speakers and 38 age and gender matched NC speakers.
- Automatic markup with manual correction:
  - a two-layer feed-forward network with sigmoid hidden and softmax output neurons identified vowel boundaries
  - the network had been trained over the spectrogram and manually produced segmentations for 1459 DDK sequences produced by PD and NC speakers
  - sealed conjugate gradient backpropagation was used as the training method
  - the preliminary markup provided by the network was processed into full V,C, and syllable segmentation by a Praat script.
- all boundaries were manually verified or adjusted

Extraction of acoustic properties of consonants, vowels, and syllables using a second Praat script.

- Computation of average consonant, vowel and syllable properties and properties of the full DDK sequence
- Prediction of
  - B) speaker group
  - B) perceived reduced or not reduced articulatory quality in read speech

- Using logistic regression with Least Absolute Shrinkage and Selection Operator (LASSO) regularization
- Evaluated outside of the training sample using 5 fold cross validation (A) and leave one out cross validation (B)

RESULTS

A: Prediction of speaker group (PD or NC)

Parameter shrinking

5 fold cross validation of logistic regression model with LASSO regularization

[pA] syllables

[kA] syllables

Reciprocal Partitioning and Regression Trees

Whole sequence properties:
- Average articulation rate
- Overall variability in syllable articulation rate
- Progressive change in articulation rate
- Short term variability in articulation rate
- Consonant to vowel amplitude ratio
- Short term variability in vowel amplitudes
- Progressive change in vowel amplitudes

Metrics identified as efficient predictors in the two investigations (A and B)

Type of metric
- Short term variability in pace [pa,ta,ka] Yes
- Short term variability in vowel amplitudes [pa,ta] Yes
- %V of syllable duration [ka] Yes
- Consonant to vowel amplitude ratio Yes
- Articulation rate [pa,ta] Yes
- Progressive change in vowel amplitudes [pa,ta,ka] Yes
- Progressive change in articulation rate [pa,ta,ka] Yes
- Prominence of release peak [pa] Yes

B: Prediction of read speech reduced articulatory quality from DDK properties

CONCLUSIONS

- Speaker group membership may be predicted with up to 93% accuracy based on an acoustic quantification of DDK sequences.
- The most accurate prediction was based on [ka] syllable sequences, which also required the fewest predictors to achieve maximum performance.
- Perception of a reduced quality of speech while reading was accurately predicted in 82% of cases based on the acoustic quantifications of DDK sequences.
- The types of metrics that were identified as strong predictors were similar overall, but
- Speaker group predictions were improved more by durational measures.
- Listeners’ assessments of articulatory quality were improved more by metrics related to vowel amplitudes and the relation to the amplitude of the consonant.