Profiling temporal patterns of connected speech in post-stroke aphasia using a semi-automated approach

Claire Cordella, PhD, CCC-SLP\textsuperscript{1,2}
Swathi Kiran, PhD, CCC-SLP\textsuperscript{1,2}

\textsuperscript{1}Center for Brain Recovery, Boston University
\textsuperscript{2}Rafik Hariri Institute for Computing and Computational Science & Engineering, Boston University

ASHA 2023
Disclosures

- C. Cordella, S. Kiran have no relevant financial or non-financial relationships to disclose
Funding

- NIH NIDCD
  - F32 DC020342 (PI: Cordella), NIH NIDCD
  - R01 DC008524 (PI: MacWhinney; AphasiaBank)
Connected speech analysis in aphasia

- Critical component of diagnostic assessment\(^1\)-\(^3\)
  - Functional, ecologically valid, and clinically feasible
  - Increasingly used in research in part thanks to large, publicly available databases of patient recordings, e.g., AphasiaBank

- Captures both what is said and how it is said
  - What: macrolinguistic, microlinguistic measures\(^4,5\)
  - How: temporal, prosodic measures\(^6,7\)

- Temporal patterns in connected speech
  - Timing of speech, as measured by detailed speech/pause measures

1. Fromm et al., *Aphasiology* (2022)
3. Fromm et al., *Semin Speech Lang* (2020)
5. Stark et al., *JSLHR* (2023)
Temporal patterns of connected speech

- Important contributor to overall fluency and communication efficiency
  - Also negatively impacts listener judgments of likeability & competence¹,²

- Reliance on coarse metrics, but mounting evidence for diagnostic potential of more detailed metrics:
  - Pause types (short, long); pause location³,⁴
  - Subcomponent speech rate measures⁵,⁶

References:
1. Harmon et al., *Aphasiology* (2016)
2. Park et al., *Aphasiology* (2011)
Current study

- Evaluate whether detailed rate/pause metrics can robustly differentiate subgroups of PWA

- Semi-automated approach to extract temporal features
  - Greater promise for clinical feasibility
  - Constrains metrics used for analysis → rate and silent pauses only
Study aims

1. Examine groups differences (correlation) across rate and silent pause measures for PWA belonging to the following categories:
   i. fluent v. non-fluent
   ii. WAB-R aphasia subtypes
   iii. WAB-R Fluency subscores

2. For groups (i)-(ii), characterize and compare cumulative silent pause duration distributions to identify differential patterns
Methods

Participants

208 persons with (post-stroke) aphasia (PWA)

3 different fluency grouping schemata

AphasiaBank recordings

Basic Demographics

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>mean age (range), yrs</td>
<td>61.8 (25.6 – 90.7)</td>
</tr>
<tr>
<td>mean YPO (range), yrs</td>
<td>5.5 (0.08 – 30)</td>
</tr>
<tr>
<td>male:female</td>
<td>118:90</td>
</tr>
<tr>
<td>mean WAB-R AQ (range)</td>
<td>71.6 (21.4 – 93.4)</td>
</tr>
</tbody>
</table>
Methods

Audio Processing Pipeline

1. Cinderella audio extracted, cross-talk excised

2. Audio noise-reduced (Audacity)

3. Custom MATLAB-based Speech Pause Analysis (SPA) software is used to automatically detect speech vs pause segments (Green et al., 2004)

4. 14 measures originally extracted following promising results in prior literature → reduced to 10 measures following correlation-based feature selection (Aim 1)

5. From SPA, (i) extracted durational values (ms) per pause per sample; (iii) log-normalized pause data; (iii) generated cumulative pause distribution densities analysis (Aim 2)

speech threshold = 25 ms
pause threshold = 200 ms
# Methods cont’d

## Temporal measures

<table>
<thead>
<tr>
<th>Measure</th>
<th>Derivation/description</th>
</tr>
</thead>
<tbody>
<tr>
<td>total duration</td>
<td>Total length (s) of sample, incl. speech + pause</td>
</tr>
<tr>
<td>speech rate*</td>
<td># words / total duration (words/s)</td>
</tr>
<tr>
<td>articulation rate*</td>
<td># words / speech duration (words/s)</td>
</tr>
<tr>
<td>percent pause</td>
<td>(pause duration / total duration)*100</td>
</tr>
<tr>
<td>mean pause duration</td>
<td>mean (individual pause event durations, s)</td>
</tr>
<tr>
<td>normalized pause count</td>
<td># pause events / total duration</td>
</tr>
<tr>
<td>mean speech duration</td>
<td>mean (individual pause event durations, s)</td>
</tr>
<tr>
<td>normalized speech count</td>
<td># speech events / total duration</td>
</tr>
<tr>
<td>coefficient of variation (cv) of pause duration</td>
<td>$cv$ (mean pause duration)</td>
</tr>
<tr>
<td>coefficient of variation (cv) of speech duration</td>
<td>$cv$ (mean speech duration)</td>
</tr>
</tbody>
</table>

*Rate variables calculated w/r/t total word count using CLAN FREQ command on paired CHAT transcripts

## Statistical Analyses

### Aim 1
- Fluent v. non-fluent (BH correction for multiple comparison)
- ANOVA, BH correction + post-hoc Tukey
- Kendall rank correlation

### Aim 2
- Kolmogorov-Smirnov (KS) tests to evaluate difference in pause distributions, correction for multiple comparison as needed
Results, Aim 1

For non-fluent aphasia, sig lower speech rate, articulation rate, mean speech duration.

Sig higher percent pause, speech and pause variability.

No sig diff in total duration, number of pause/speech events, or mean pause duration.
Results, Aim 1

All seven vars sig differentiate Broca’s aphasia from more fluent subtypes (Conduction, Anomic, Wernicke's)

Speech rate and percent pause show most robust differentiation across all subgroups
Results, Aim 1

Sig **positive** corr b/w speech rate, articulation rate, mean speech duration and WAB Fluency subscore

Sig **negative** corr for percent pause, variability of speech and pause duration, and mean pause duration
Results, Aim 2

Sig different cumulative pause duration distributions for...

fluent v. non-fluent aphasia (p<.001)

Short and long pauses common in fluent aphasia
Long pauses more common in non-fluent aphasia

Broca’s v. Wernicke’s | Anomic | Conduction (p<.001)

Anomic v. Wernicke’s | Conduction (p<.01)
Discussion

- Feasibility of semi-automated temporal measures to differentiate fluency profiles in post-stroke PWA
  - Overall, **important variables for differentiation were consistent** no matter which way ‘fluency’ was spliced
  - Joins emerging work on dx utility of temporal measures in post-stroke aphasia\(^1,2\)

- Future direction(s) to improve clinical viability
  - Streamlining of audio processing pipeline to minimize manual labor
  - Pairing of temporal metrics with linguistic features for fuller picture of non-fluent aphasia

---

**References**

1. DeDe, Salis, AJSLP (2020)
THANK YOU!

cordella@bu.edu

Acknowledgments:

Lauren Di Filippo, Natalie Spitalnic, James Rice, Kamryn Schult (*audio processing, analysis*)

AphasiaBank database (*all primary data*)

Jordan Green, Brian Richburg (*SPA MATLAB code*)
Selected References