Dimensions of spontaneous speech in aphasia: A factor analysis

(*JSLHR*, in press)

JEAN K. GORDON, PHD, CCC-SLP
COMMUNICATION SCIENCES & DISORDERS
UNIVERSITY OF IOWA

- ***partially supported by funding from the ASH Foundation (New Century Scholars Grant, 2017)***
- ***presented to the Academy of Aphasia, online, October 18-20, 2020***
- ***currently in press in JSLHR***

2. Motivation

- ***Spontaneous speech tasks*** *are critically important for the clinical diagnosis of aphasia, both because of their* ***linguistic complexity*** *and their* ***ecological validity****.*
- *However, their linguistic complexity also makes performance* ***highly variable****, both across and within aphasia subtypes.*
- *Aphasia syndromes are criticized for their variability—is there a* ***better basis for classification?***

The current study identified factors underlying spontaneous speech in aphasia, and investigated their relationship to traditionally defined aphasia subtypes via linear discriminant and latent profile analyses.

3. 19 microlinguistic variables…

- Narrative-level measures:
  1) total utterances; 2) speech rate (WpM)
- Utterance-level measures:
  3) MLU; 4) sentence complexity;
  5) embeddings; 6) grammatical errors;
  7) repairs; 8) circumlocution/empty speech;
  9) jargon
- Word-level measures:
  10) MATTR; 11) content:function ratio;
  12) propositional density; 13) verb marking;
  14) *noun marking**; 15) morphological errors;
  16) *semantic errors**; 17) phonological errors;
  18) unrelated word errors; 19) neologisms

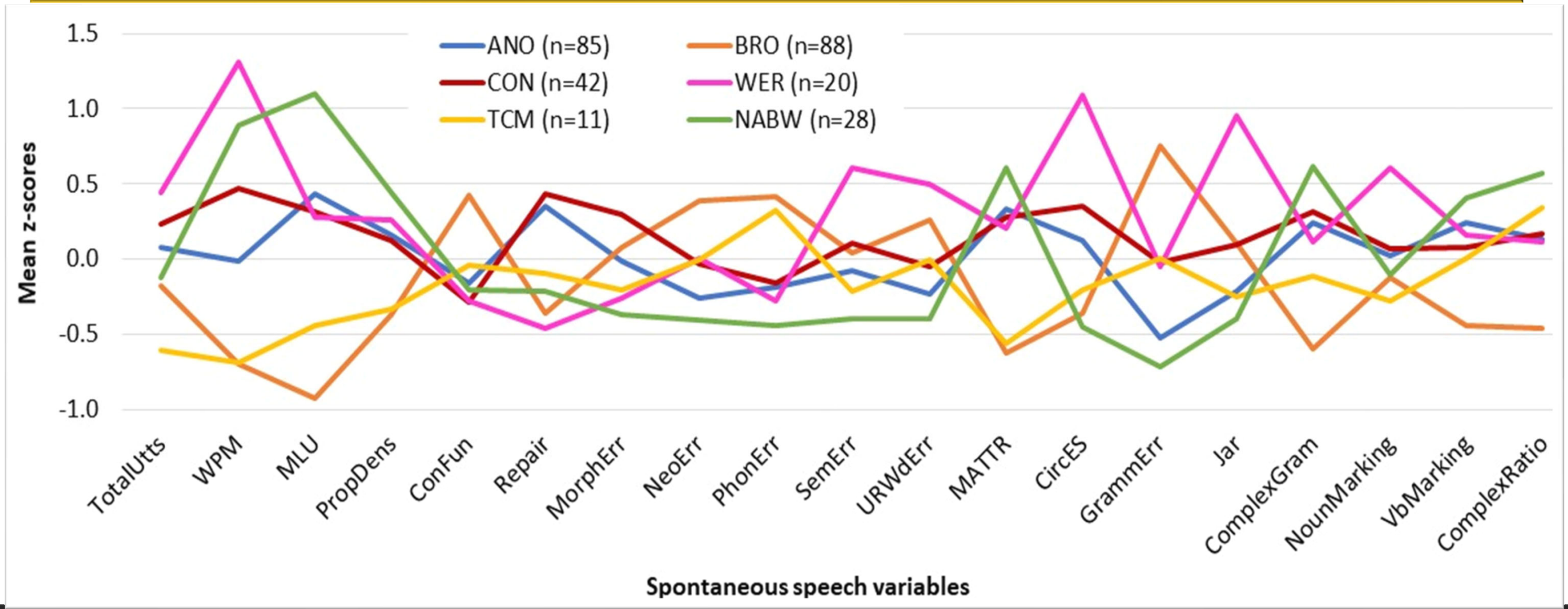
…in 274 persons with aphasia (PwA)…

- Aphasia Quotients: 13–99 (median=75)
- Types of aphasia:
  - 88 Broca's aphasia (32%)
  - 85 anomic aphasia (31%)
  - 42 conduction aphasia (15%)
  - 20 Wernicke's aphasia (7%)
  - 11 transcortical motor aphasia (4%)
  - 28 not aphasia by WAB, or NABW (10%)

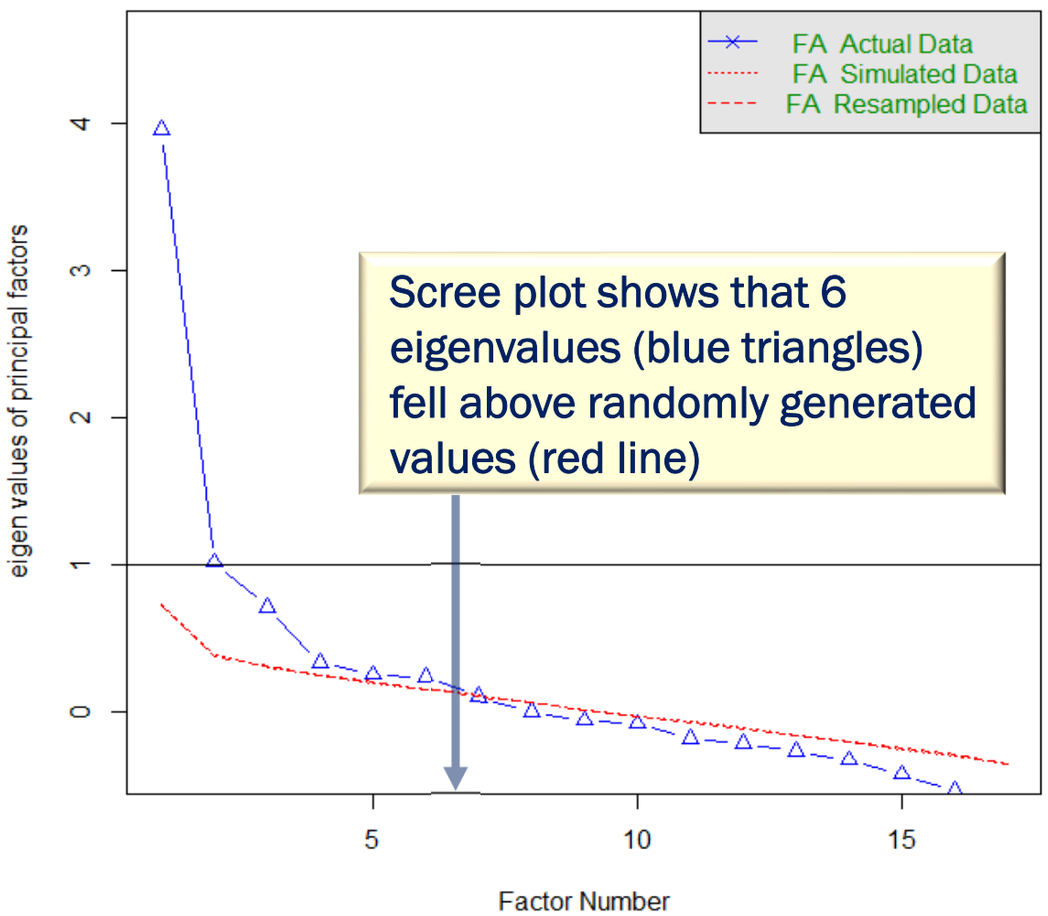
…analyzed from story retelling task (Cinderella)

* Two variables were removed because of a lack of shared variance, indicating poor 'factorability'.

4. Mean performance by syndrome showed expected patterns, especially on measures of fluency (WpM, MLU), grammatical complexity & accuracy, circumlocution & jargon



5. Exploratory factor analysis using oblique rotation (*promax*) generated 6 factors:

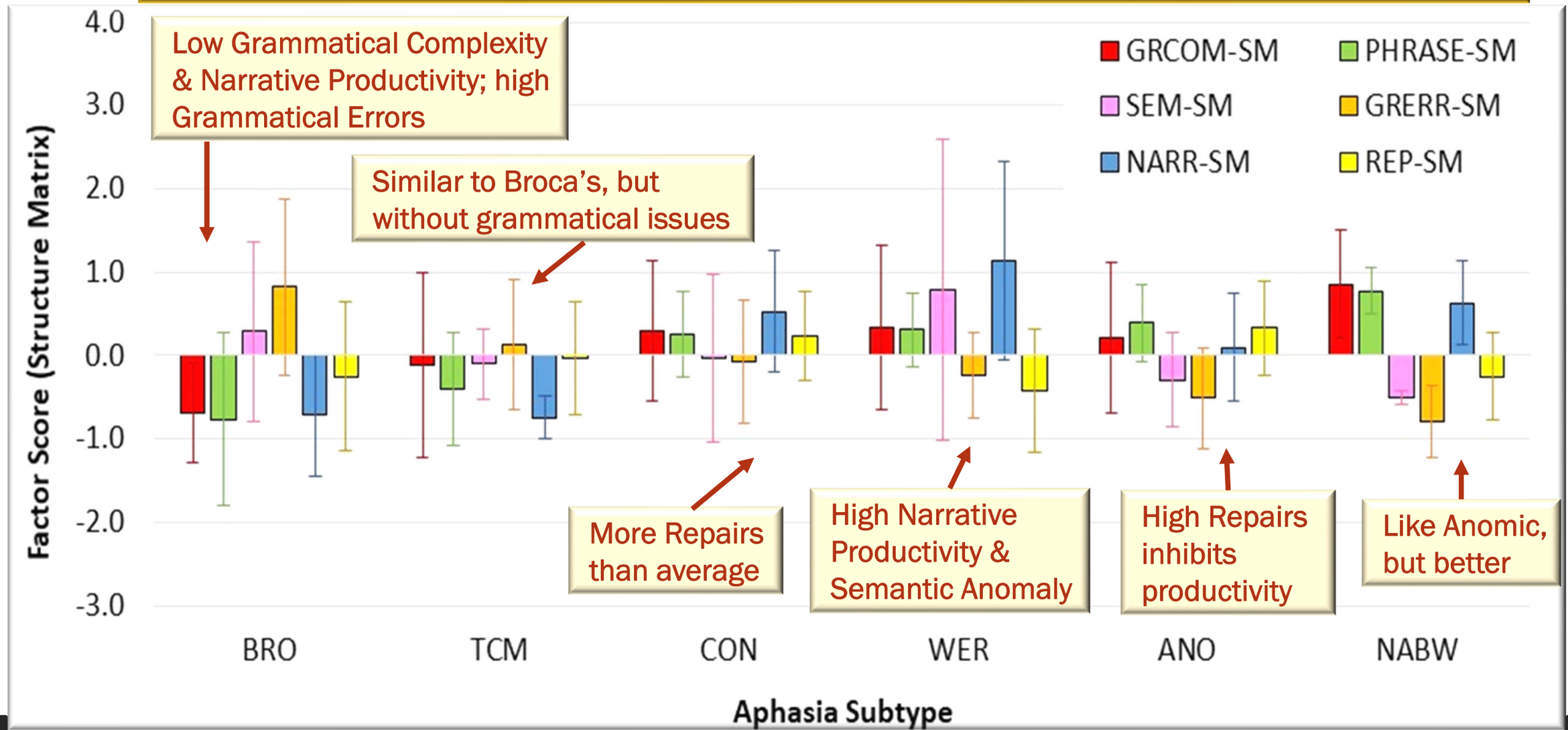


Factors	Unique Variance (pattern matrix)	Total Variance (structure matrix)
Phrase Building (ability to form basic word groupings to form propositions)	13 %	19 %
Grammatical Complexity (proportion of syntactically complex structures)	10 %	18 %
Semantic Anomaly (production of semantically incorrect utterances)	10 %	13 %
Grammatical Errors (production of syntactically incorrect utterances)	8 %	14 %
Narrative Productivity (production of more speech, more quickly)	7 %	13 %
Repairs (production interrupted by repetitions and retracing)	4 %	5 %
Total	*52 %*	*NA*

6. Factor structure matrix (loadings > .4 are highlighted)

Variable	Phrase Building	Grammatical Complexity	Semantic Anomaly	Grammatical Errors	Narrative Productivity	Repairs
Total Utterances	0.10	-0.02	-0.04	-0.08	0.37	0.14
Speech Rate (WpM)	**0.47**	**0.59**	-0.15	-0.39	**0.86**	-0.14
Utterance length (MLU)	**0.69**	**0.79**	-0.32	**-0.55**	**0.53**	0.20
Propositional Density	**0.70**	0.33	-0.31	-0.16	0.29	0.10
Content:Function Ratio	**-0.44**	-0.13	0.02	**0.46**	**-0.41**	-0.26
Repairs (repeat, retrace)	0.25	0.20	-0.17	-0.14	0.20	**0.58**
Morphological Errors	-0.12	-0.12	-0.06	0.28	-0.02	0.20
Neologistic Errors	**-0.52**	-0.24	**0.57**	0.27	-0.31	-0.25
Phonological Errors	**-0.45**	-0.25	0.14	0.33	-0.38	-0.19
Unrelated Word Errors	-0.31	-0.20	**0.68**	0.06	-0.23	-0.28
Type-Token Ratio (MATTR)	**0.61**	**0.43**	-0.31	-0.36	0.30	0.07
Circumlocution/Empty Speech	0.19	0.24	-0.05	-0.19	0.34	0.14
Grammatical Errors	**-0.41**	-0.33	0.02	**0.97**	-0.37	-0.15
Jargon	-0.17	-0.15	**0.96**	-0.11	0.03	-0.02
Embedded Clauses	**0.49**	**0.86**	-0.25	-0.34	0.35	0.21
Verb Marking	**0.48**	0.18	-0.12	-0.30	0.16	0.18
Complexity Ratio	0.31	**0.82**	-0.17	-0.25	0.18	0.07

7. Factor scores were generated for each individual (*regression method*) and compared across syndromes



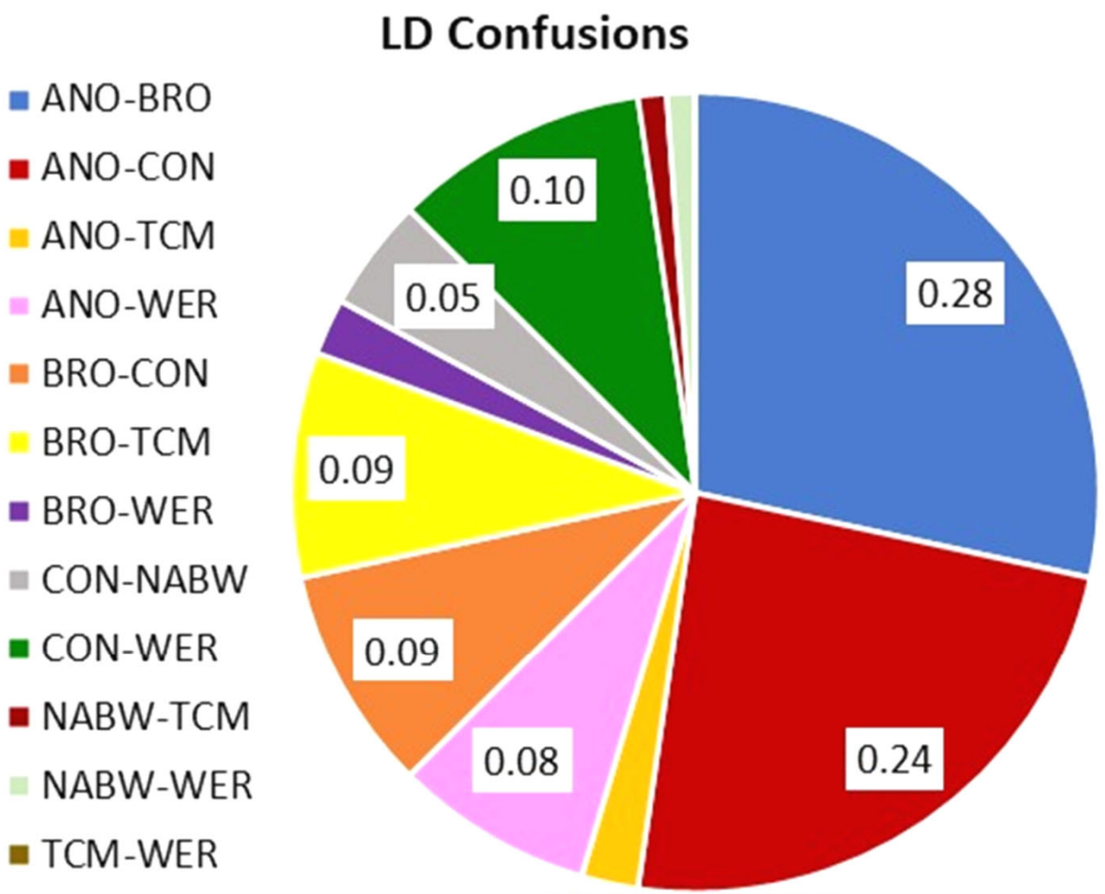
8. Linear discriminants were used to calculate the probability of each speaker belonging to their given aphasia subtype

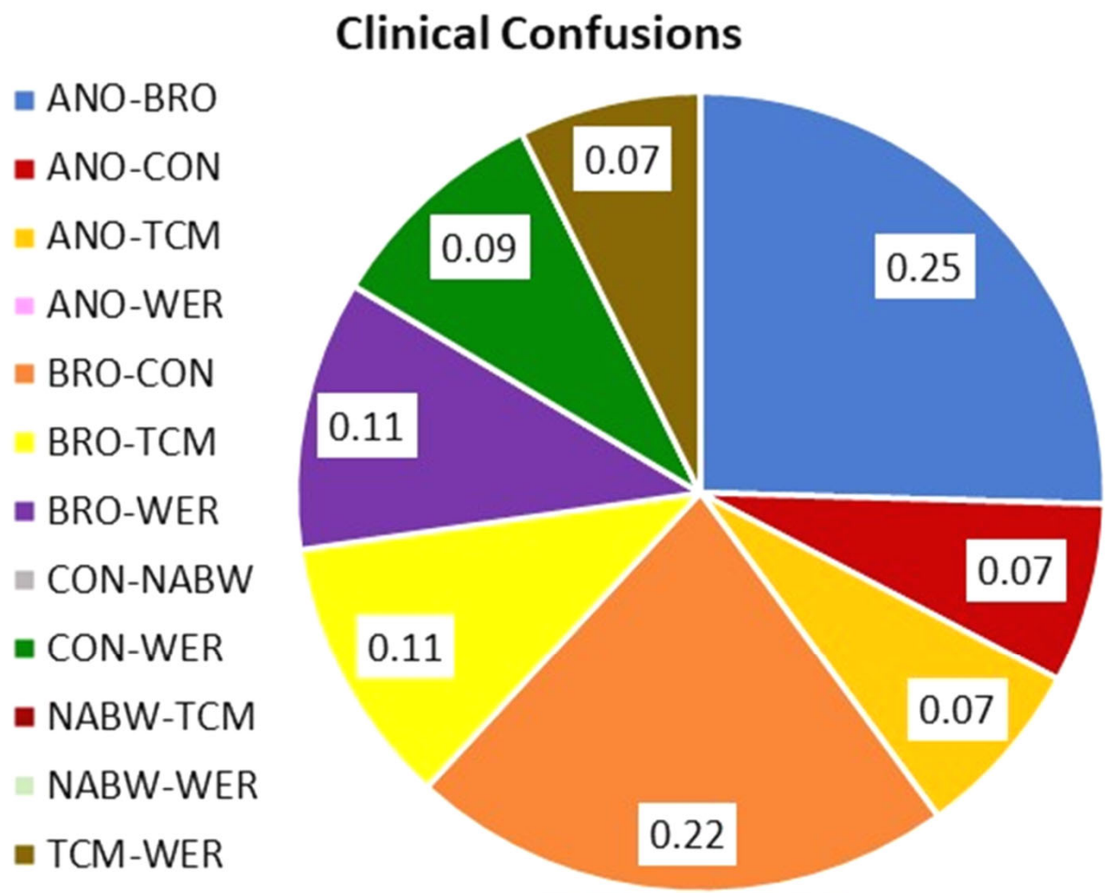
Number (%) of each subtype correctly classified by linear discriminants

Diagnosed subtype	BRO	TCM	CON	WER	ANO	NABW	Predicted Sums
Broca's aphasia	68 (0.77)	8 (0.73)	6 (0.14)	1 (0.05)	8 (0.09)	0 (0.00)	91 (1.03)
Transcortical motor aphasia	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)
Conduction aphasia	2 (0.02)	0 (0.00)	10 (0.24)	4 (0.20)	3 (0.04)	1 (0.04)	20 (0.48)
Wernicke's aphasia	1 (0.01)	0 (0.00)	5 (0.12)	11 (0.55)	4 (0.05)	0 (0.00)	21 (1.05)
Anomic aphasia	17 (0.19)	2 (0.18)	18 (0.43)	3 (0.15)	66 (0.78)	16 (0.57)	122 (1.44)
Not aphasic by WAB	0 (0.00)	1 (0.09)	3 (0.07)	1 (0.05)	4 (0.05)	11 (0.39)	20 (0.71)
Actual sums	88 (1.00)	11 (1.00)	42 (1.00)	20 (1.00)	85 (1.00)	28 (1.00)	166 (0.61)

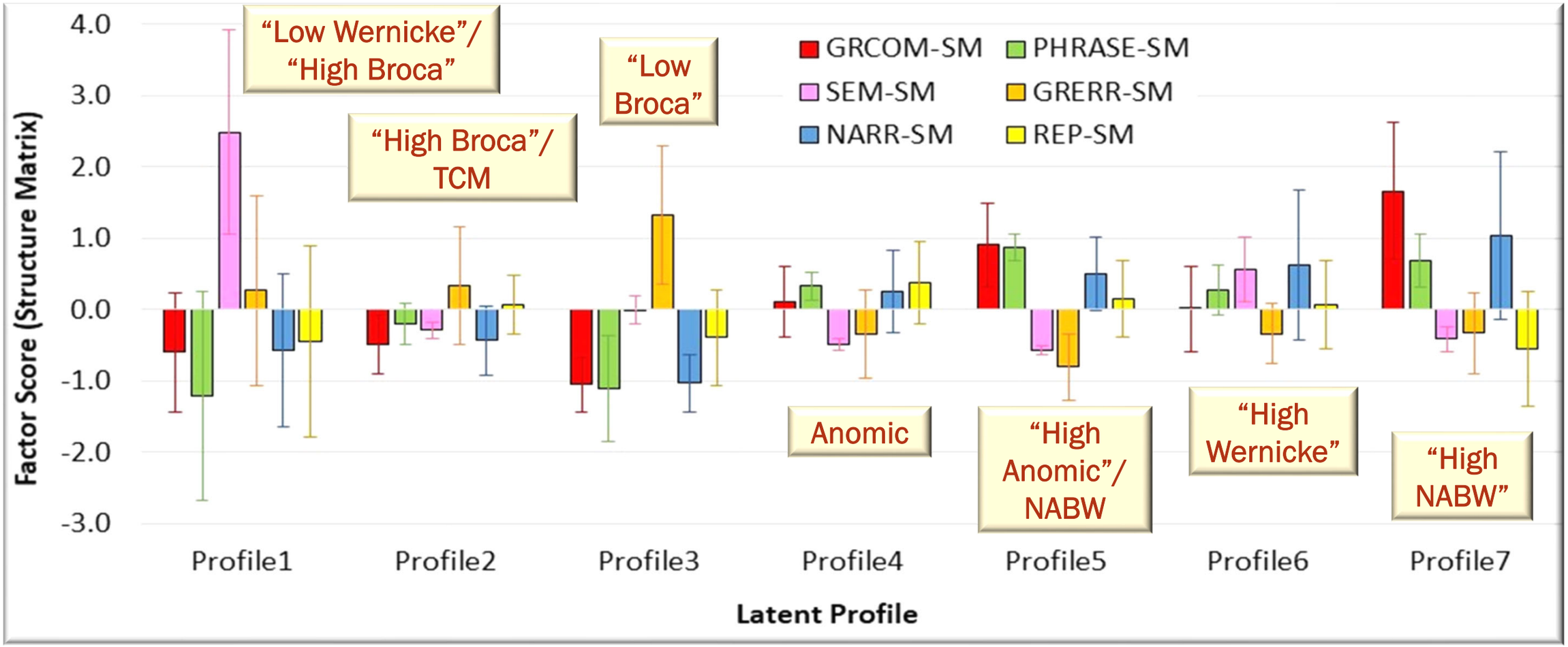
- Largest groups (Broca's, anomic) were more accurate, but over-predicted.
- Lowest accuracy for TCM (0%) and conduction aphasia (24%).
- Overall, 61% were correctly classified.

9. Both the LD model and clinicians were prone to confuse anomic and Broca's aphasia. The model was more likely to confuse conduction with anomic aphasia; clinicians were more likely to confuse conduction with Broca's aphasia.



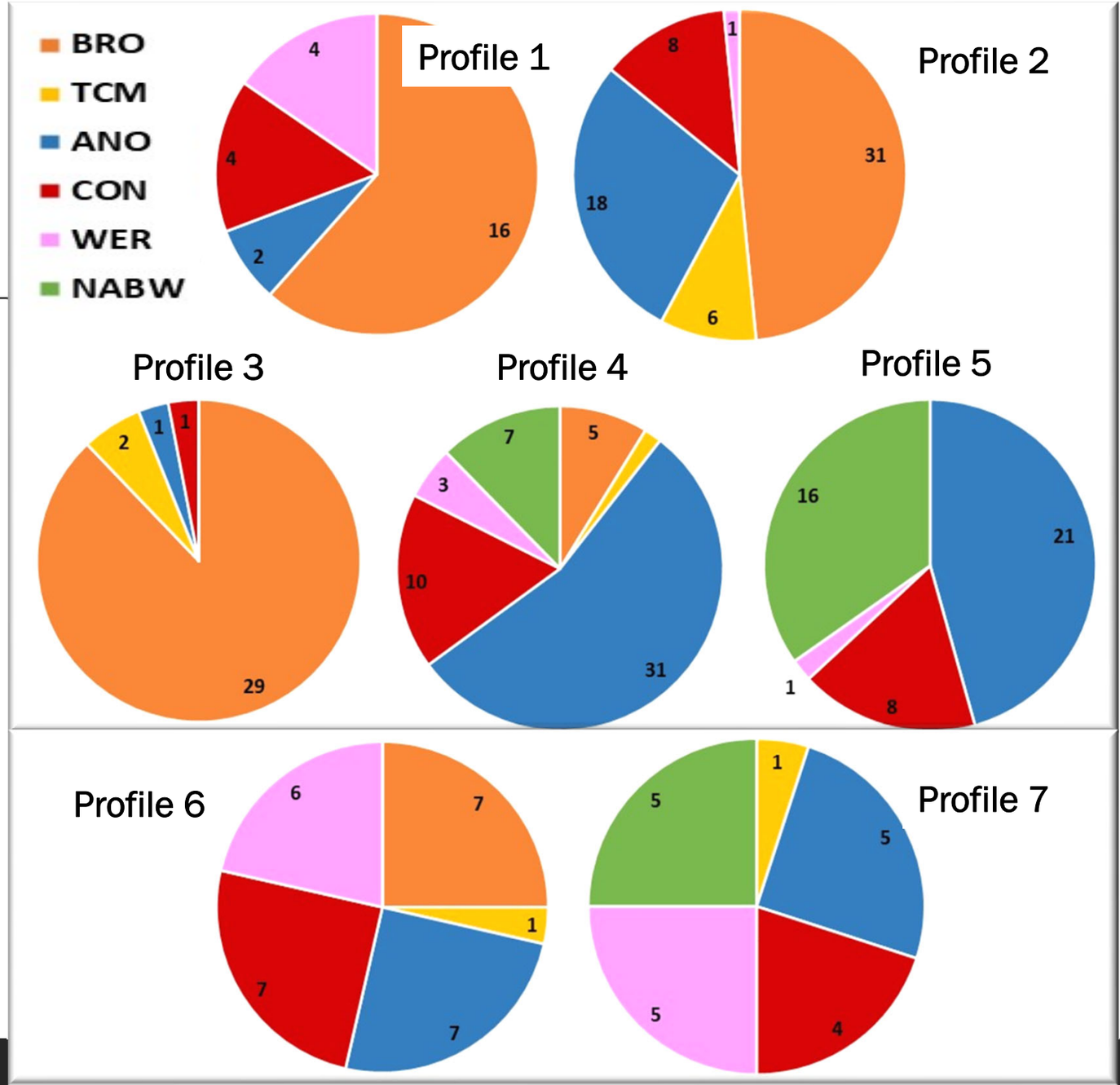


10. A latent profile analysis of the factor scores generated 7 profiles that captured qualitative differences among syndromes and quantitative differences within syndromes.



11. Latent profiles showed little correspondence to subtype diagnoses.

- Each profile included at least 4 aphasia subtypes
  - even Profile 4 (“Anomic”) included 5 other subtypes
- Each aphasia subtype was represented in at least 3 of the profiles.
- Performance differences were confirmed between ‘high’ and ‘low’ categories on variables and AQ scores.



12. Summary & Discussion

- *An **exploratory factor analysis** generated **6 factors** reflecting phrase-level and narrative-level fluency, grammatical accuracy and complexity, semantic accuracy, and repair behaviors.*
- *Factors accounted for **52% of the variance**, leaving open questions about other sources of variance (e.g. premorbid style).*
- ***Linear discriminant analyses** generated correct subtype classifications for only ~60% of the speakers. Mismatches were similar to those shown by clinicians.*
- ***Latent profile analyses** reflected both quantitative and qualitative variance in generating classifications, dividing many syndromes into higher and lower-performing subsets.*

- *Limitations reinforce the need to **balance** such data-driven, **population-level approaches** with more fine-grained, **individual-level analyses**.*