

## Lexical diversity

- the most widely used measure of verbal skills
- is known to be lower in discourses by people with aphasia (PWA) as compared to healthy speakers (Fergadiotis & Wright, 2011)
- Current study — apply various lexical diversity measures to speech samples from Russian speakers with and without aphasia

## Lexical diversity measures

### Type-token ratio (TTR)

number of words in a sample divided by the number of lexemes

- + easy to calculate
- very dependent on sample size

### Measure of textual lexical diversity (MTLD)

mean length of sequential word strings in a text that maintain a given TTR value (0.72)

- + independent of sample size

### Moving average type-token ratio (MATTR)

average of TTR measured in a window of n:  
word 1 - word n, word 2 - word (n+1), ...

- window size 10 - can detect such properties of the text, as frequent repetitions
- window size 100 - not sensitive to repetitions

- + independent of sample size

(Covington & McFall, 2010)

## Russian CiPS corpus

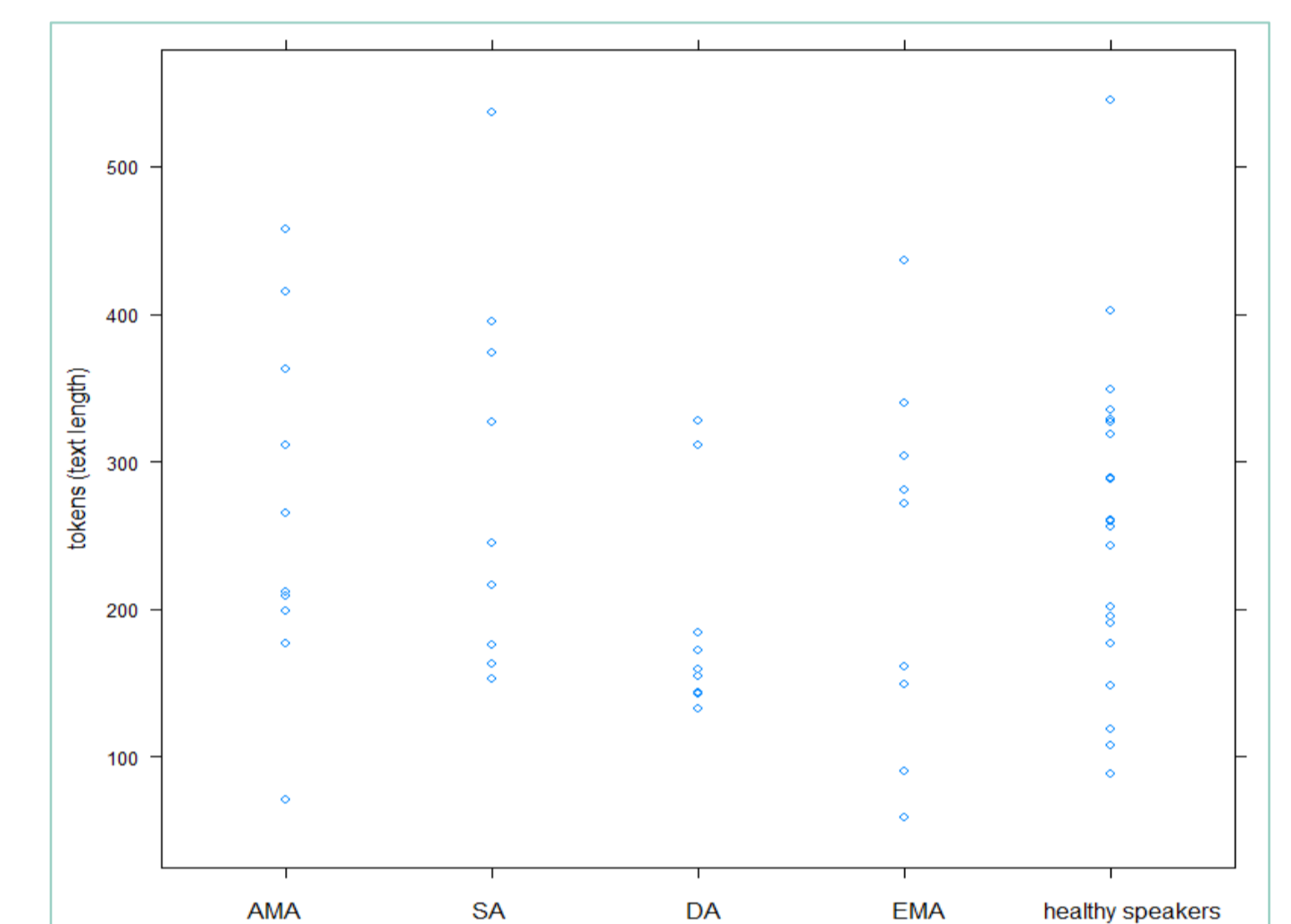
- oral retellings of the Pear film (Chafe, 1980; Khudyakova et al., 2016)
- comparable samples and clear story line
- annotation on multiple levels, including lexical

## Texts

Group		N	Age (mean; range)
Fluent	Acoustic-mnemonic aphasia (AMA)	10 (5 females)	51.3; 40-68
	Sensory aphasia (SA)	9 (4 females)	58.3; 33-81
Non-fluent	Dynamic aphasia (DA)	9 (5 females)	51.8; 41-68
	Efferent motor aphasia (EMA)	10 (3 females)	48.6; 30-57
Healthy speakers		21 (10 females)	58.4; 25-84

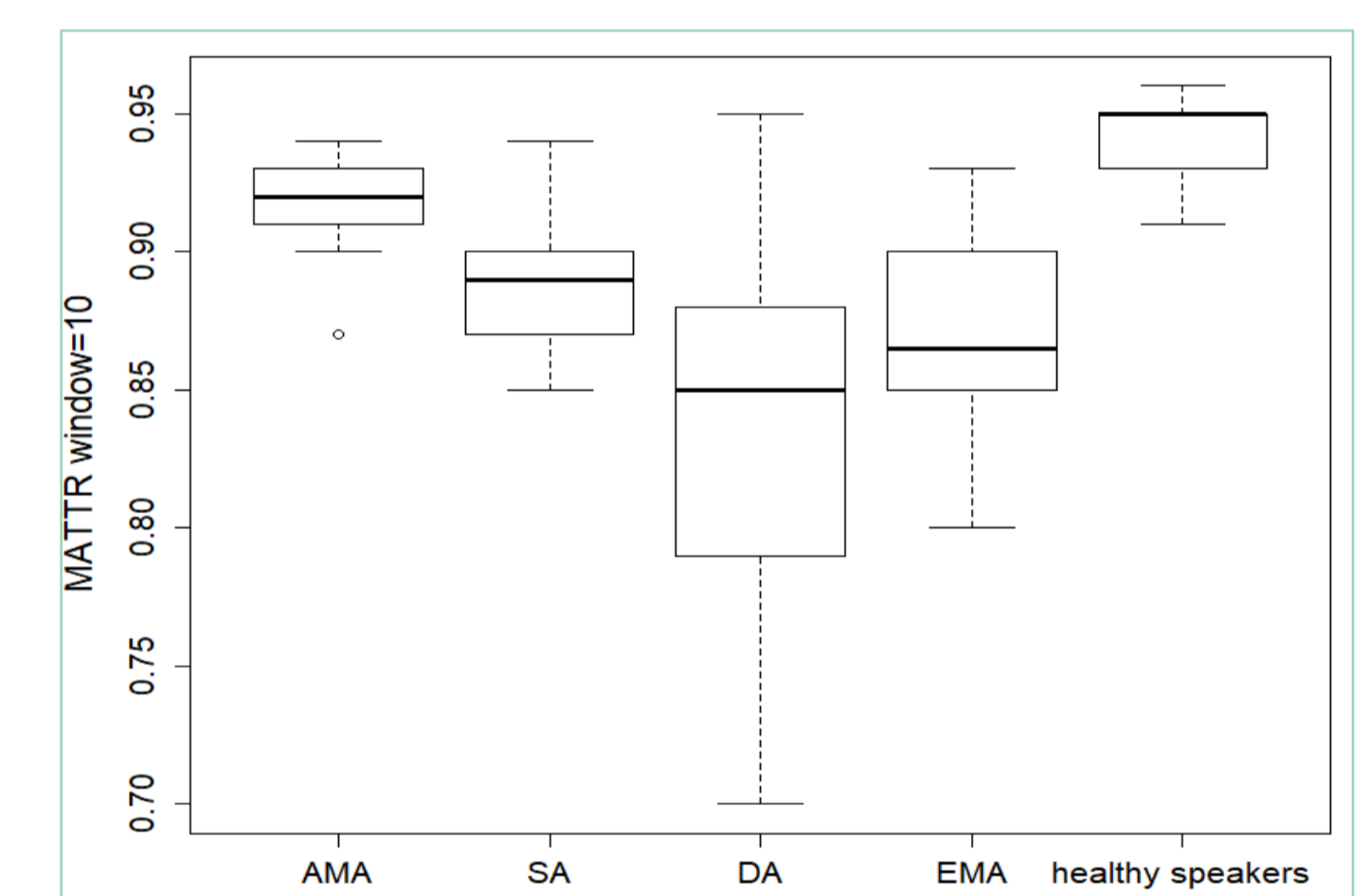
## Results

Great variability in text length



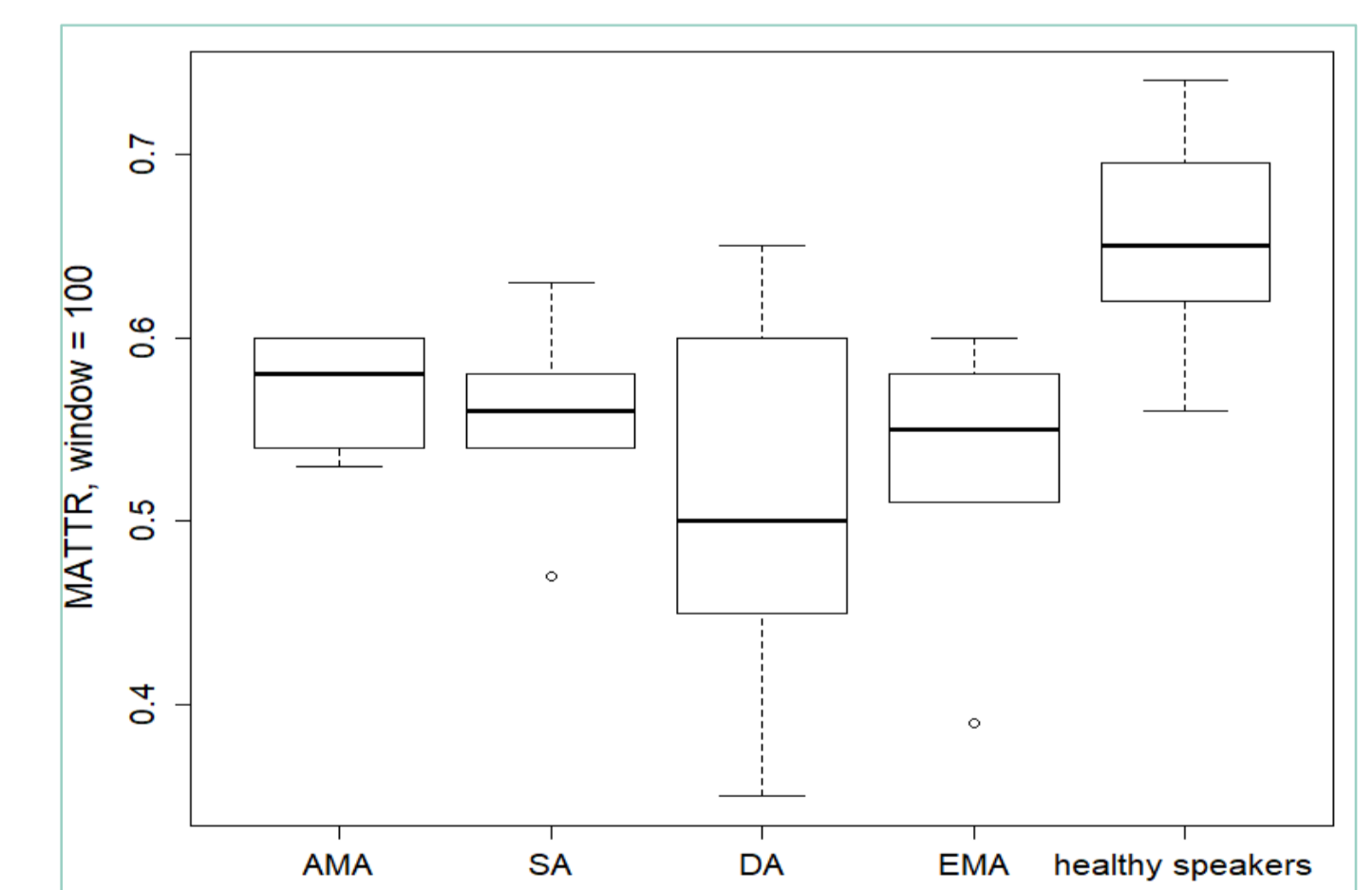
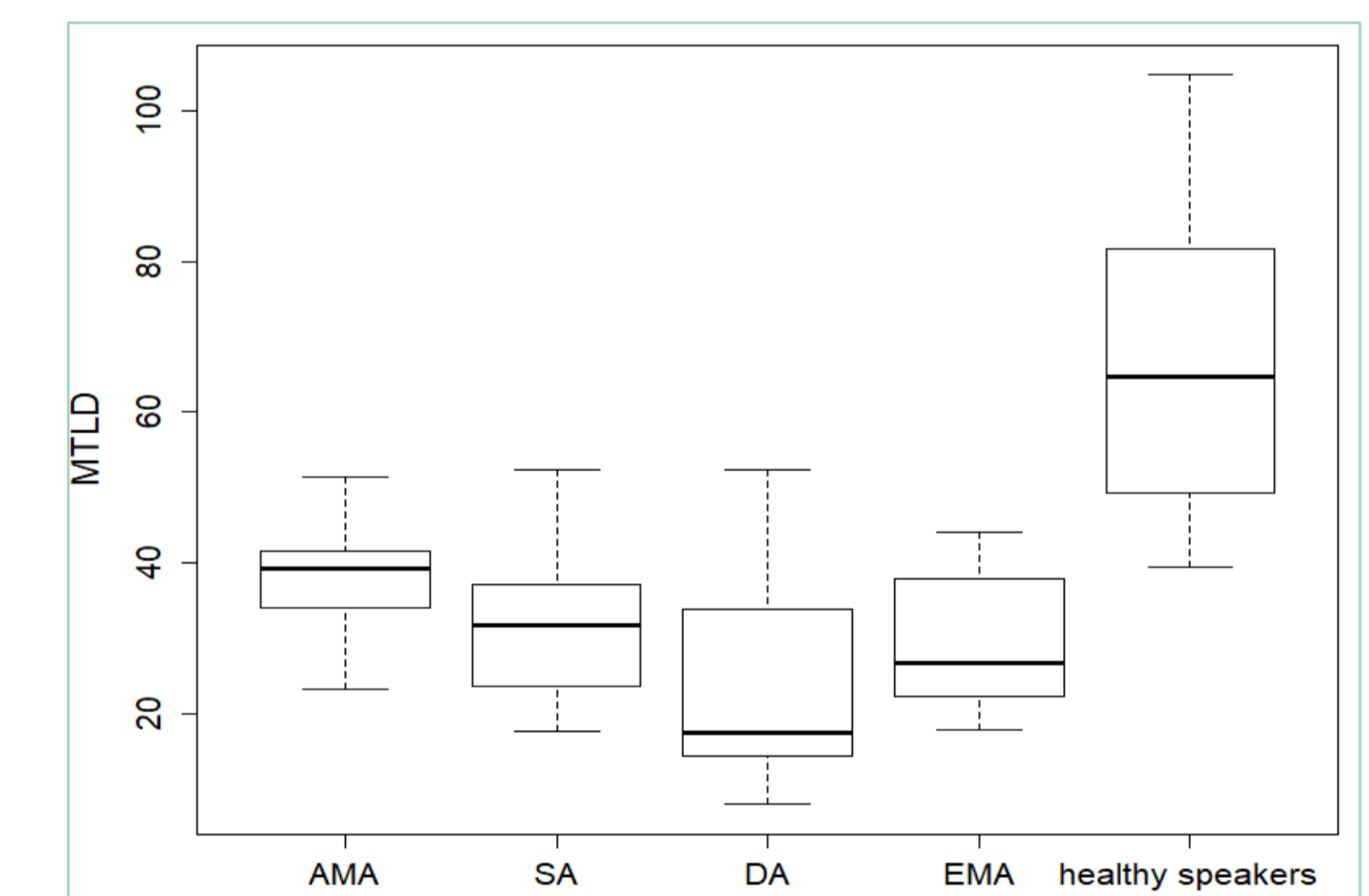
### MATTR with window =10

- Significant difference between healthy speakers and all PWA groups
- Significant difference between speakers with dynamic aphasia and speakers with fluent types of aphasia (sensory and acoustic-mnemonic)



### MTLD and MATTR with window=100

- Significant difference between healthy speakers and all PWA groups
- No significant difference between speakers with different types of aphasia



## Discussion

PWA have lower lexical diversity scores than healthy speakers: replicates previous findings

A specific measure of lexical diversity – MATTR with a smaller window size can detect differences between texts by speakers with dynamic aphasia and fluent aphasia types.

## References

- Chafe, W. (1980). *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*. (W. Chafe, Ed.). Norwood, New Jersey: Ablex.
- Covington, M. a., & McFall, J. D. (2010). Cutting the Gordian Knot: The Moving-Average Type–Token Ratio (MATTR). *Journal of Quantitative Linguistics*, 17(2), 94–100.
- Fergadiotis, G., Wright, H., & West, T. (2013). Measuring lexical diversity in narrative discourse of people with aphasia. *American Journal of Speech-Language ...*, 22(May), 397–409.
- Khudyakova, M. V., Bergelson, M. B., Akinina, Y. S., Iskra, E. V., Toldova, S., & Dragoy, O. V. (2016). Russian CiPS: a Corpus of Narratives by Brain-Damaged Individuals. In *LREC Proceedings*. Portoroz, Slovenia.