Distribution of Language Measures among Individuals with and without Non-Fluent Aphasia

Paula Garcia
Rochester Institute of Technology
1 Lomb Memorial Drive
Rochester, NY, USA
pxg5962@rit.edu

ABSTRACT

Existing studies report varied findings on the features that influence a non-fluent aphasia diagnosis, due to the numerous measurements and aphasia classifications available. This study applies Mann-Whitney U test to evaluate the distribution of 32 language measures among 78 individuals with and without non-fluent aphasia in the *AphasiaBank* corpus. The results highlight the different distribution of verbs, auxiliaries, and present participles between the two populations. Future studies will combine the results of this study with the measures speech therapist deem most informative, to generate a tool for visually synthesizing the *AphasiaBank* data.

CCS Concepts

•**Human-centered computing** → *Accessibility theory, concepts and paradigms;*

Keywords

Mann-Whitney U; Aphasia; Language; Test battery

1. INTRODUCTION

Approximately one third of stroke survivors are diagnosed with aphasia and more than 1 million people in the United States are currently living with aphasia [4]. Aphasia diagnoses are generally split into fluent and non-fluent aphasia, with the latter defined by effortful communication of limited syntax but conserved comprehension.

The primary methods for aphasia diagnosis are formal and informal testing, which include results from clinician elicited conversation samples, Western Aphasia Battery and Boston Diagnostic Aphasia Examination. Often, these diagnostic methods yield contradictory results [2], which is why supplemental measures are used to ascertain therapy goals and methods [1]. Prior studies have evaluated the measures that contribute to an aphasia diagnosis using the Pearson's chi-square [1] and Welch's t-test [3], yielding significant results for mean length per utterance, syntax errors, and word errors. The extent to which multiple measures contribute to a final diagnosis is debated due to incomparable studies with small sample sizes, varied sampling criteria, and limited evaluated measures [6].

The purpose of this study is to identify the population measures whose medians significantly differentiate a non-fluent aphasia outcome, as identified by either the Western Aphasia Battery, Boston Diagnostic Aphasia Examination, or clinician evaluations. The *AphasiaBank* corpus [5] is purposefully used in this study to support comparative findings as the population sample grows. For each language measure (n=32), the distributions of individuals with (n=39) and without non-fluent aphasia (n=39) are compared:

H_0: *The distribution of values for individuals with and without non-fluent aphasia are equal.*
H_1: *The distribution of values for individuals with and without non-fluent aphasia are not equal.*

The results of this study will be presented to *AphasiaBank* users to inform them of sampling limitations and the measures that differ between both populations.

2. METHODOLOGY

The data was sourced from the *AphasiaBank* corpus which includes comparable language measures of discourse and clinical tests. The sampled populations with (N_{orig}=204) and without (N_{orig}=202) non-fluent aphasia were pre-processed using R, yielding 39 individuals with non-fluent aphasia and 63 without non-fluent aphasia. As outlined in Table 1, individuals who had multiple etiologies or were missing complete datasets (discourse or test results) were excluded from the sample. Next, individuals with (Mdn.$_{age}$=75, IQR=11) and without (Mdn.$_{age}$=63, IQR=20.24) non-fluent aphasia were matched, as best as possible, by ascending age due to the significant variance between both populations (Levene's test, median adjusted df, α=0.05: F(1,86.03)=4.66, p=0.034).

All 32 common language measures were tested for homogeneity of variance (Levene's test) and the statistically significant differences of their distributions (Mann-Whitney U) using SPSS. Non-parametric tests were used due to the presence of non-normal distributions.

3. RESULTS

The Levene's test for homogeneity of variance (median, adjusted df) yielded significant results (α=0.05) for type-token ratio, verbs per utterance, word errors, utterance errors, idea density, 3rd person singular, past tense verbs, past

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

PETRA '17 June 21-23, 2017, Island of Rhodes, Greece

© 2017 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-5227-7/17/06.

DOI: http://dx.doi.org/10.1145/3056540.3076214

Table 1: Preprocessing for populations with and without non-fluent aphasia (NFA) in *AphasiaBank*

Exclusion criteria	Sequentially updated sample size: **NFA** (n=204)	**Without NFA** (n=202)
Diagnosis of non-fluent aphasia	204	-
No multiple observations	152	193
No missing data	131	193
Monolingual	114	162
No apraxia	42	-
No dysarthria	40	-
Protocol followed	39	63

Table 2: Significant results for Levene's test for homogeneity of variance (α=0.05)

Measure	Result
Type-token ratio	$F(1,58.90)=6.73$, $p=0.012$
Verbs per utt.	$F(1,69.99)=4.54$, $p=0.037$
Word errors	$F(1,38.06)=24.80$, $p<0.001$
Utt. errors	$F(1,38.89)=32.66$, $p<0.001$
Idea density	$F(1,39.66)=38.74$, $p<0.001$
3rd person sing.	$F(1,71.66)=8.35$, $p=0.005$
Past tense	$F(1,98.07)=4.80$, $p=0.031$
Past participle	$F(1,61.45)=4.17$, $p=0.046$
Conjunction	$F(1,58.19)=6.70$, $p=0.012$
Retracing	$F(1,40.02)=5.01$, $p=0.031$
Repetition	$F(1,39.88)=6.24$, $p=0.017$

participles, conjunctions, retracing, and repetition, as shown in Table 2. The remaining 21 out of 32 language measures were used to assess the differences in distribution among the two populations using the Mann-Whitney U test.

Three out of the 21 measures had non-significant (α=0.05) differences in distribution among the two groups: mean length per utterance (MLU) utterances (U=583.50, p=0.077), total utterances (U=610.50, p=0.134), and duration (U= 864.50, p=0.299). All remaining language measures are ranked in order of significance, with ascending p-values all less than 0.001: verbs (U=167.50), auxiliaries (U=212), identical forms of 1st and 3rd person singular (U=260.5), past tense verbs (U=203), present participle (U=220.5), prepositions (U= 113), adjectives (U=237), adverbs (U=234), determiners (U=213.5), MLU words (U=10), MLU morphemes (U=7), types (U=69), tokens (U=176), words per minute (U=10), morphemes per word (U=176), nouns (U=166), pronouns (U=214.5), and plural words (U=248).

4. DISCUSSION & CONCLUSION

Eleven out of 32 measures could not be tested for significance in distribution due to an already underlying difference in the variance of distributions. The number of word errors for instance, had zero occurrences among individuals without aphasia and multiple occurrences among individuals with non-fluent aphasia. The context when these measures occur among each population may be more informative than the comparison between populations.

The remaining 21 out of 32 measures were tested for significance. Only 3 out of 21 measures supported the null hypothesis, indicating that the discourse samples had similar distributions by MLU utterances, total utterances, and duration. A similar distribution of utterances and duration supports the comparison of discourse samples. The number of words (tokens) however, had significant differences between both populations.

Eighteen out of 21 measures had significant results to support the alternate hypothesis, *The distribution of values for individuals with and without non-fluent aphasia are not equal.* All 18 statistically significant measures had a p-value less than 0.001, highlighting the consistent differences between the two populations. The distribution of verb, auxiliaries, identical forms of 1st and 3rd person singular counts were the top 3 measures in regards to statistical significance.

Future studies will combine the measures deemed most important by speech language therapists, with the measures identified in this study. The combination of these measures will be used to develop a tool for speech language therapists to review and compare the data in *AphasiaBank*.

5. REFERENCES

[1] C. Code and D. Rowley. Age and aphasia type: The interaction of sex, time since onset and handedness. *Aphasiology*, 1(4):339–345, 1987.

[2] M. A. Crary, R. T. Wertz, and J. L. Deal. Classifying aphasias: Cluster analysis of western aphasia battery and boston diagnostic aphasia examination results. *Aphasiology*, 6(1):29–36, 1992.

[3] K. C. Fraser, F. Rudzicz, and E. Rochon. Using text and acoustic features to diagnose progressive aphasia and its subtypes. In *INTERSPEECH*, pages 2177–2181. Citeseer, 2013.

[4] J. Galliers, S. Wilson, A. Roper, N. Cocks, J. Marshall, S. Muscroft, and T. Pring. Words are not enough: empowering people with aphasia in the design process. In *Proceedings of the 12th Participatory Design Conference: Research Papers-Volume 1*, pages 51–60. ACM, 2012.

[5] B. MacWhinney, D. Fromm, M. Forbes, and A. Holland. Aphasiabank: Methods for studying discourse. *Aphasiology*, 25(11):1286–1307, 2011.

[6] S. O'Connor and C. M. Pettigrew. The barriers perceived to prevent the successful implementation of evidence-based practice by speech and language therapists. *International journal of language & communication disorders*, 44(6):1018–1035, 2009.