Brian MacWhinney
Carnegie Mellon University

Davida Fromm
Carnegie Mellon University

Margie Forbes
Carnegie Mellon University

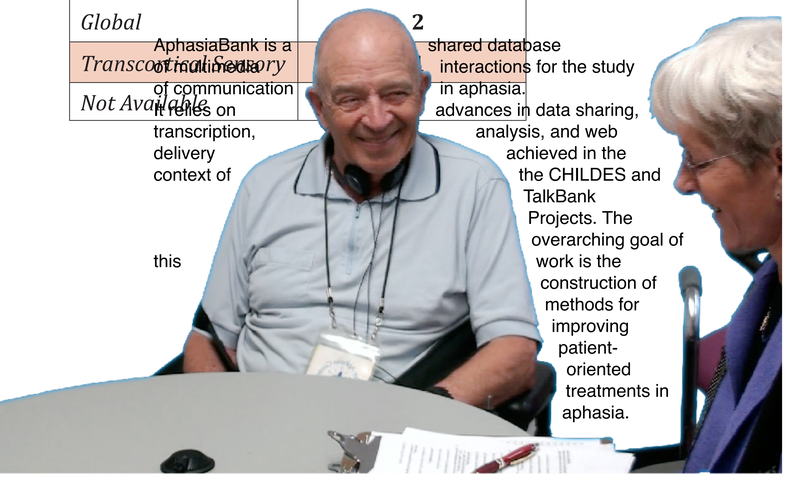
Audrey Holland
University of Arizona

Comprehensive Statistical Analysis of Aphasia Type

Anomic	33
Broca	25
Conduction	14
Wernicke	9
Above WAB Cutoff	9
Transcortical Motor	5
Global	2
Transcortical Sensory	
Not Available	

What is AphasiaBank?

AphasiaBank is a shared database of multimedia interactions for the study of communication in aphasia. It relies on advances in data sharing, transcription, analysis, and web delivery achieved in the context of the CHILDES and TalkBank Projects. The overarching goal of this work is the construction of methods for improving patient-oriented treatments in aphasia.



PURPOSE

Purpose: Classification is important for theory and practice.
Does therapy depend on type or only severity?
Do types reflect neural processes?
Do types reflect recovery stages?

Issue: Traditional measures (WAB, Aachen, BNT) do not study conversational or discourse abilities

Question: Can we devise new measures that include a fundamental role for conversational and discourse abilities?

Question: Can we do this within the framework of a publicly shared database -- AphasiaBank

Question: How will the results of this analysis align with traditional measures?

Question: How will these results align with clinician judgments?

METHOD

METHOD:

1. **Population**. Samples and scores for this analysis come from the AphasiaBank database that now provides transcriptions of test sessions with 176 PWAs, along with standardized test scores.
2. **Discourse Protocol**. Participants from 14 different sites were tested with a standardized protocol,
   a. free speech (stroke story and recovery)
   b. three picture descriptions
   c. story telling (Cinderella)
   d. procedural discourse (making a peanut butter and jelly sandwich).
3. **Non-coversational measures**. We also administered the Aphasia Quotient subtests from the Western Aphasia Battery-Revised (WAB; Kertesz, 2007), the short form Boston Naming Test-Second Edition (Kaplan, Goodglass, & Weintraub, 2001), the Northwestern Verb Naming Test (Thompson, in preparation), and a non-standardized test developed to assess word-level and sentence-level repetition.
4. **Video Recording**. All discourse tasks and testing were recorded on video.
5. **Demographics**. Extensive demographic data (51 fields) were collected on each participant as well.

TRANSCRIPT ANALYSIS:
1. Transcription was done using the CHAT format, which operates closely with the CLAN programs (MacWhinney, 2000).
2. We coded word repetitions, revisions, fillers, gestures, sound fragments, and unintelligible output. Word-level errors were coded in six categories: phonology, semantics, neologism, dysfluency, morphology, and formal lexical features.
3. The CLAN programs produced automatic morphosyntactic analysis of the transcripts from which we calculated total words, total utterances, duration of PWA response, and vocabulary diversity (VOCD, Malvern et al. 2005).

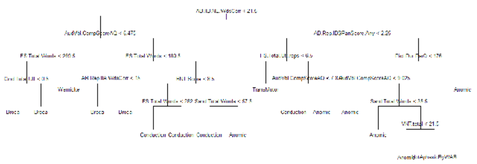
TRANSCRIPTION

```
*PAR:  &=points:firemen &=imit:siren .
*INV:  &=laughs .
*PAR:  yeah &=points:cat &=points:man .
*INV:  do you think it's gonna [: going to] work out ?
*INV:  do you think he [//] that's gonna [: going to] +/?
*PAR:  +< oh yeah &=head:yes &=points:firemen .
*INV:  +< yeah what do you think is gonna [: going
       to] happen ?
```

Statistical Analysis

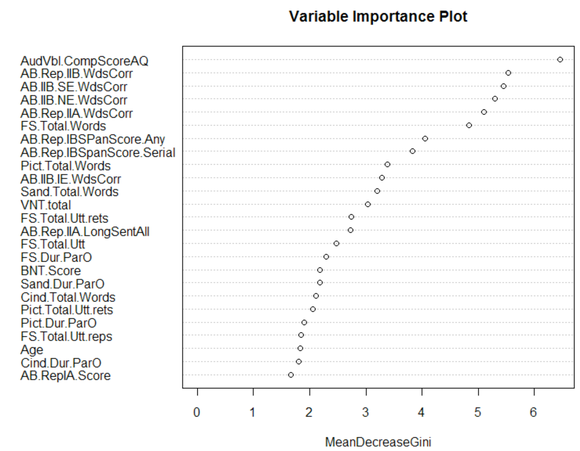
Statistical Analysis:

1. **Variable Reduction**. The first steps involved reduction of the data set to drop out highly correlated variables. We used Principal Components analysis to pull out a set 25 variables from the original 134.
2. **Participant Clustering**. We then used k-means clustering to extract the most meaningful set of participant groups. This was done first by maximizing the Adjusted Rand Index (ARI) and the Fowlkes-Mallows Index (FMI).
3. **WAB vs Conversation**. At this point it was clear that the WAB could be used to determine clusters, but we wanted to consider what clusters would form without relying on the WAB. However, we left in the auditory comprehension score from the WAB.
4. **Random Forest**. We then computed a random forest of decision trees to evaluate classification. From this we could evaluate the Gini coefficient of each measure. (see next column).
5. **Five Groups**. From the random forest analysis we extracted a pattern of five groups that could be characterized in terms of their position on these four dimensions: repetition, naming, comprehension, and fluency. These groupings only partly overlap with WAB characterizations as shown in the figure in the RESULTS section.
6. **Model-based Clustering**. We also used model-based clustering that maximized the BIC (Bayesian Information Criterion) with three clusters and a VEI (Variable Shape, Equal Volume, Diagonal Orientation) density. This method showed that participants classified as Broca by clinicians fell into two different clusters in the statistical analysis, as suggested earlier by Sundet and Engvik (1985), and that those classified as Anomic also split into two separable types.

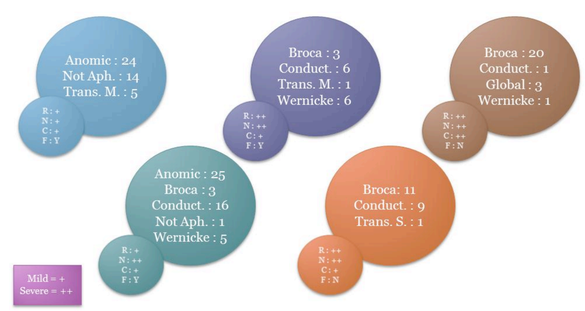


VARIABLE ANALYSIS

Type to enter text



RESULTS - 5 groups



We can characterize these 5 proposed classification types in order of severity, in this way:

1. High Functioning (light blue): This group primarily includes PWAs classified by the WAB as Anomic and non-aphasic (i.e., their WAB scores were above its cutoff, but still have demonstrable language difficulties).

2. Anomic (green): These PWAs largely have difficulties in naming, but include individuals who were classified as having Conduction aphasia. These individuals demonstrate only moderately impaired repetition skills.

3. Non-classifiable (purple): This group is mixed by WAB standards, but can be characterized as fluent with moderate comprehension, and with severe naming and repetition problems. Note that this group includes the majority of our participants who were judged as Wernicke on the WAB along with an equal number of individuals classified as Conduction on that test.

4. Broca with good comprehension (tan): These PWAs display problems in fluency with adequate comprehension. This group includes participants who are characterized as having Conduction aphasia on the WAB, but who differ from those in groups 2 and 3 by displaying more substantial problems in repetition.

Name That Aphasia

Participant				WAB
1	B	B	B	B
2	B	NC	TM	A
3	B	NC	B	B
4	B	TM	A	A
5	W	W	W	W
6	A	W	TS	W
7	B	B	B	B
8	W	A	C	C
9	B	NC	B	C
10	B	NC	TM	B
11	NC	B	C	W
12	B	B	B	A
13	B	B	TM	B
14	G	G	B	G
15	W	C	NC	W
16	B	B	B	C
17	NC	TM	A	A
18	C	B	C	C
19	G	G	B	B
20	B	B	A	B
21	C	C	C	C
22	W	W	W	W
23	G	G	G	G
24	NC	W	C	W
25	A	A	A	A
26	B	B	NC	B
27	A	A	C	C
28	A	A	A	A
29	B	B	B	B
30	TM	NC	NC	B
29				B
27				

DISCUSSION

This study opens the way to a more clinically-relevant approach as well as one that is consistent with newer, growing understanding of the nature of language processing and its breakdown in aphasia.

By anallyzing our four basic conversational measurements, along with repetition and comprehension tests, a new aphasic patient from outside of our study may be categorized into one of the five proposed groups below. We next plan to use the scheme to classify a new sample of 40 PWAs. If these participants can be assigned cleanly to groups, the scheme will be validated.

Analyses of this type can help us modernize the traditional classifications based on assumptions of lesion locations. As we learn more about how the brain processes language, we may need to reformulate our traditional conceptions of aphasia disorders.

PROBLEMS:

Underrepresentation of PW Global Aphasia.
Underrepresentation of rarer groups, including Wernicke.
Need for conversationally relevant comprehension measures.

This project is funded by NIH_NIDCD grant RO1-DC008524 (2007-2017)