

Language structure across dementia types and changes during disease progression

Vitor Zimmerer¹, Chris Hardy², Rebecca Bond², Lucy Russell², Jonathan Rohrer², Jason Warren², Rosemary Varley¹



¹Department of Language and Cognition, University College London

²Dementia Research Centre, University College London



Background

- Language production changes over course of dementia.
- Changes have a direct impact on the ability to participate in family and society, and therefore influence wellbeing and quality of life.
- In Alzheimer's disease, language analysis has been sensitive to disease progression (Zimmerer, Wibrow, & Varley, 2016).

Aims

- Use of an automated analysis to examine language production across different dementia types.
- Characterize features of words and word combinations.

Language phenomena relevant to this study

Content words: Refer to things (*table, election*), actions (*find, smash*) or characteristics (*blue, slow*). Marker of lexical-semantic impairment.

Function words: Express grammatical relationships, e.g., pronouns (*I, they, her*), articles (*the, a*), question words (*why, where*). Marker of grammatical impairment.

Usage frequency: Indicates how often words or word combinations are encountered in everyday communication. More frequent units are easier to produce. More frequent combinations are also easier and may become "fossilized" (formulaic language). Higher frequency units are more resilient to neural degeneration.

Fluency: The degree to which words are produced in combinations or in isolation. Reduced verbal fluency is referred to as "telegraphic".

Methods

- Secondary analysis of data collected by the UCL Dementia Research Centre.
 - Semantic variant primary progressive aphasia (svPPA)
 - Logopenic variant primary progressive aphasia (lvPPA)
 - Non-fluent variant primary progressive aphasia (nfPPA)
 - Behavioural variant frontotemporal dementia (bvFTD)
 - Neurotypical controls (NC)
- Elicitation of spontaneous connected speech: "Can you tell me about your last holiday?"
- Samples were transcribed for computerized analysis with the FLAT.

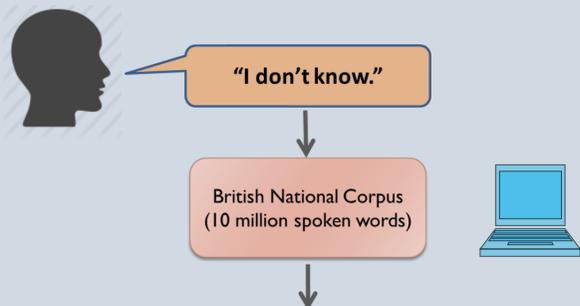
Group	n	Age (M)	MMSE (M)
svPPA	29	64	20.4
lvPPA	26	63.3	19.3
nfPPA	34	69.9	25
bvFTD	14	64.4	23.8
NC	20	62.8	-

Frequency in Language Analysis Tool (FLAT)

Developed at UCL (Zimmerer, Wibrow, & Coleman, 2017)

Works with any orthographic transcription.

Extracts words, bigrams (two-word combinations) and trigrams (three-word combinations) from the transcript and extracts their usage frequencies and collocation strengths (t-scores) from the British National Corpus (BNC).



	I	do	n't	know
Word frequency	30956	9956	12628	5743
Bigram frequency	-	2157	4222	1065
Bigram t-score	-	125.9	199.3	96
Trigram frequency	-	-	1866	904
Trigram t-score	-	-	136.3	95

Categorizes words as content or function words.

Computes "combination ratio" as an indicator of fluency: $\frac{n \text{ of trigrams}}{n \text{ of words}}$

Results

Between-group comparisons

- Multinomial logistic regressions
 - Group as DV
 - All IVs residualized over Age
- NC differed from all other groups in three IVs:
 - Combination ratio, $\chi^2 = 99.848, p < .001$.
 - Content word frequency, $\chi^2 = 40.741, p < .001$.
 - Collocation strength, $\chi^2 = 18.472, p = .001$.
- Pairwise comparisons (n.s. = blue; $p < .05$ = yellow; $p < .01$ = orange; $p < .001$ = red)

	svPPA	lvPPA	nfPPA	bvFTD
NC	Yellow	Red	Red	Red
svPPA	Blue	Yellow	Yellow	Yellow
lvPPA	Blue	Blue	Blue	Blue
nfPPA	Blue	Blue	Blue	Blue

	svPPA	lvPPA	nfPPA	bvFTD
NC	Yellow	Red	Blue	Yellow
svPPA	Blue	Yellow	Blue	Blue
lvPPA	Blue	Blue	Blue	Blue
nfPPA	Blue	Blue	Blue	Blue

	svPPA	lvPPA	nfPPA	bvFTD
NC	Red	Red	Red	Yellow
svPPA	Blue	Yellow	Blue	Blue
lvPPA	Blue	Blue	Blue	Blue
nfPPA	Blue	Blue	Blue	Blue

	svPPA	lvPPA	nfPPA	bvFTD
NC	Yellow	Yellow	Yellow	Yellow
svPPA	Blue	Blue	Blue	Blue
lvPPA	Blue	Blue	Blue	Blue
nfPPA	Blue	Blue	Blue	Blue

Marker of disease progression?

- Correlations between language values and Age and MMSE scores, within each group.
 - Pearson's r, one-tailed.
- No sig. correlations with Age.
- MMSE and Combination ratio:
 - lvPPA, $r = .432, p = .037$.
 - bvFTD, $r = .673, p = .008$.
- MMSE and Content word frequency:
 - svPPA, $r = -.593, p = .001$.
 - lvPPA, $r = -.488, p = .02$.
- MMSE and collocation strength:
 - bvFTD, $r = -.587, p = .022$.

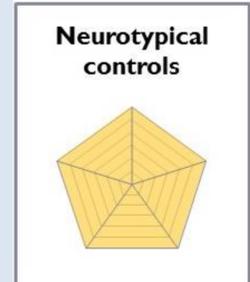
Summary and discussion

- Most prominent features of dementia were reduced fluency as well as use of more common content words and word combinations (formulaic language).
- All profiles differed from another, and fluency and content word properties were sufficient to distinguish all groups from another.
- Variables show some sensitivity to general cognitive impairment.
- Tools like FLAT show potential in assisting diagnosis and tracking change over time, including as the result of intervention.

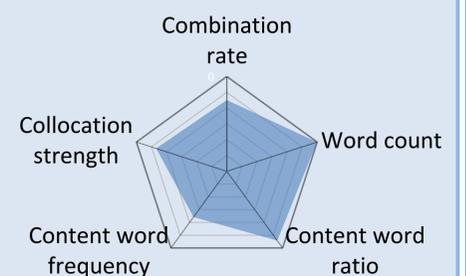
References

Zimmerer, V. C., Wibrow, M., & Varley, R. A. (2016). Formulaic language in people with probable Alzheimer's Disease: a frequency-based approach. *Journal of Alzheimer's Disease*, 53, 1145–1160.
Zimmerer, V. C., Wibrow, M., & Coleman, M. (2017). Frequency in Language Analysis Tool.

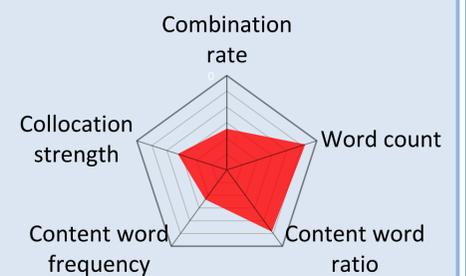
Language profiles (normalized to NC variance)



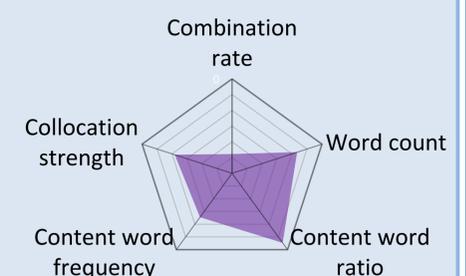
Semantic variant PPA



Logopenic variant PPA



Non-fluent variant PPA



Behavioural variant FTD

