SABER 2021

Applying Ecological Diversity Methods to Improve Examination of Constructed Responses

Megan Shiroda

Michael Fleming & Kevin Haudek

Constructed Response reveals student thinking.

- Constructed response assessment requires students to use their own words to explain a complex topic.
- For instructors: provides valuable information on student understanding.
- For researchers:
 - Reveals student thinking and unique language.
 - How does thinking and language change?

Your friend loses 15lbs on a diet. Where did the mass go?

DEVELOPING IDEAS

- General Metabolism
- Mass Converted to Energy
- Excretion
- How to Lose Weight

SCIENTIFIC IDEAS

- Correct Products
- Exhalation
- Molecular Mechanism

"Most of the mass that the friend lost is transferred to <u>exhalation</u> of **carbon dioxide** through the lungs. The left over mass is excreted in *urine, sweat, breath, and tears.*"

Data set has categorical data.



Thinking Type - Based on the inclusion/exclusion of ideas

Scientific – only scientific ideas

Mixed – combination of scientific and developing

Developing – only developing ideas

Worked heavily with this data set: Uhl et al. 2021; Shiroda et al. *in review*

Lexical diversity doesn't measure diversity in CRs.

LEXICAL DIVERSITY

- Traditionally used to assess student essay writing and language proficiency.
- Ratio of unique words to total words (Type to Token, TTR)
- Difficult to apply to constructed response due to length
 - ~22.5 words per response... Concatenate?
- Difficult to interpret
 - How repetitive are the responses?
 - Not, how do they differ from each other?
 - What does low vs high TTR really mean?
- The big cat ate the bird. TTR=5/6
 The bird is sitting on a cat. TTR=7/7 TTR=9/18
 The cat ate the bird. TTR=4/5

Ecological diversity measures could be used instead.

ECOLOGICAL DIVERSITY

- Examines and compares the species in an environment or environments.
- Not limited by length / size
- Techniques are quantitative, easily interpretable and have a visual output.

ECOLOGICAL	Species 1	Species 2	 Species <i>n</i>
Site 1	35	10	 0
Site 2	20	0	 0
Site <i>n</i>	3	18	 56

	LEXICAL	Word 1	Word 2	 Word <i>n</i>	
	Response1	1	0	 1	7
	Response 2	0	1	 1	
s				 	/,
	Response <i>n</i>	0	1	 0	

418 response 254 words

Hypothesis:

- Ecological techniques will better quantify lexical diversity.
 - Biggest difference in Thinking > Timing > Type
- Visualization methods will assist in examining a corpus of responses

This work is confirmatory; however, we envision the techniques to be used in an exploratory manner for novel data sets.

For the data set or within groups:



For the data set or within groups:

- Whittaker's β-diversity
 - Compares between two responses
 - How many unique words?



For the data set or within groups:

- Whittaker's β-diversity
 - Compares between two responses
 - How many unique words?
- Species Turnover (Half-changes)
 - "Half change" compares the total number of species found in two or more sites
 - 50% shared, 50% unique = 1 half change
 - After 4 half-changes, essentially share no species.



Ecological measures are more interpretable than lexical diversity measures.

TTR

- Appears that the largest difference is in **Types**, but this was not observed in qualitative analysis.
 - **RICU** responses are <u>longer</u>, which would would affect TTR.

	Whole	r	Thinking			Timing		Institutional Type		
	set	Dev	Mixed	Sci	Pre	Post	TYC	PUI	RICU	
Type to Token	0.0262	0.1373	0.1385	0.1277	0.112	0.1029	0.1229	0.1013	0.1762	
Beta Diversity	37.4	37.4	31	57.3	35.2	39.9	39.3	37.7	35.5	
Half changes	2.3	2.3	2	2	2.4	2.2	2.4	2.4	2.2	

Ecological measures are more interpretable than lexical diversity measures.

Beta diversity

- Most distinct difference based on Thinking.
 - Scientific responses have a higher proportion of unique words.
 - Could be a reflection of length of responses generally shorter.

	Whole '		Fhinking		Timing		Institutional Type		
	set	Dev	Mixed	Sci	Pre	Post	TYC	PUI	RICU
Type to Token	0.0262	0.1373	0.1385	0.1277	0.112	0.1029	0.1229	0.1013	0.1762
Beta Diversity	37.4	37.4	31	57.3	35.2	39.9	39.3	37.7	35.5
Half changes	2.3	2.3	2	2	2.4	2.2	2.4	2.4	2.2

Ecological measures are more interpretable than lexical diversity measures.

Half changes

- \sim 2 half changes means about a 25% of words are shared.
- Lowest with Mixed and Scientific responses indicates Mixed and Scientific responses are more similar to themselves.

	Whole		Thinking		Timing		Institutional Type		
	set	Dev	Mixed	Sci	Pre	Post	TYC	PUI	RICU
Type to Token	0.0262	0.1373	0.1385	0.1277	0.112	0.1029	0.1229	0.1013	0.1762
Beta Diversity	37.4	37.4	31	57.3	35.2	39.9	39.3	37.7	35.5
Half changes	2.3	2.3	2	2	2.4	2.2	2.4	2.4	2.2

Also want to examine differences between groups.

- Ordination methods
 - Data reduction technique
 - Compares responses based on the inclusion of words.
 - Each response is represented as a point.
 - Similarity is shown by how close or far apart two points are.
 - Categorical data can be overlayed.
- Many ordination methods choose based on the data and intentions.
- Detrended Correspondence Analysis
 - The first axis is units of half changes.
 - 100 units = 1 half change



Ordination provides a visual depiction of the corpus





Ordination allows comparison among groups.



Ordination allows comparison among groups.



Ordination allows comparison among groups.



Statistical Tests: PERMANOVA

Compares all points in a given group to all the points in another group to determine if the groups are distinct.



PERMANOVA allows statistical comparison among groups.



Conclusions

- Ecological diversity measures are more interpretable for analyzing lexical diversity of CRs.
- Ordination provides a visual representation of the corpus.
- Overlaying categorical data allows for easy interpretation of differences between groups in the categories.
- PERMANOVA provides statistical strength to claims of categorical differences.

Acknowledgements

AACR Group: Michael Fleming, Kevin Haudek

Leonora Kaldaras, Juli Uhl, Jennifer Kaplan

Kamali Sripathi, Jenifer Saldanha, Lauren Jescovitch, Marisol Mercado Santiago, Xiaoming Zhai, Matt Steele, John Merrill, Mark Urban-Lurain

Contact Information:

Megan Shiroda

shirodam@msu.edu

beyondmultiplechoice.org











This material is based upon work supported by the National Science Foundation (DUE grants 1323162). Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the supporting agencies.