

Productive Vocabulary across Discourse Types

Introduction

According to Phillips and Jorgensen (2002), "...what underlies the word "discourse" is the general idea that language is structured according to different patterns that people's utterances follow when they take part in different domains of social life"(p. 1). This definition emphasizes the predictable nature of structural similarities among discourse instantiations, based on which discourse is categorized in different kinds, or *genres*.

It is generally accepted that genres impose different cognitive and linguistic demands on a speaker (Bliss & McCabe, 2006). As a result, the performance on several indices used to assess the microstructure and macrostructure of discourse production can vary significantly depending on the genre. Further, performance on some indices can also vary as a function of age, due to the change in cognitive functioning associated with normal aging.

Given the well documented clinical and research utility of discourse, it is very important to establish how the nature of each discourse type influences the sampling process. This knowledge allows clinicians and researchers to maximize the potential of discourse as a medium to explore human communication in aging as well as in neurologically impaired adults. Further, it enhances the ability to make comparisons across different types of discourse.

Though some aspects of discourse production have attracted considerable interest (e.g., information units), one of the most illuminative predictors of oral performance, *lexical diversity* (LD), is often neglected or, when used, it is often estimated incorrectly. Tools that had been available for comparing LD across different genres or participant groups included type/token ratio (TTR), various algebraic transformations of TTR¹, and probabilistic models such as Yule's K (Yule, 1944). These measures are notorious for covarying with sample length, thus yielding mathematically and conceptually spurious results (Malvern & Richards, 1997; Tweedie & Baayen, 1998; Jarvis, 2002; Vermeer et al., 2000)

Recently a new measure, *D*, has been developed that combines an algebraic transformation model and curve fitting. *D* is significantly robust to length variation thus allowing comparisons of different genre samples within and between participants as well as across studies (McKee et al., 2000).

The purpose of this study is two-fold: (1) to explore whether there are differences in LD among four types of discourse (procedural discourse, single picture descriptions, story telling, and recounts) and (2) to assess to what extent age influences LD when using each of the aforementioned types of discourse.

Method

Participants

A total of 86 adults with no known neurological impairments participated, 43 younger (YG) and 43 older (OD) adults. The mean ages for the groups were 23.00 years (*SD* =

¹ For, example, Split TTR (Engber, 1995), Root TTR (Guiraud, 1960), Corrected TTR (Carrol, 1964), Log TTR Herdan (1960)

1.98) and 75.28 years ($SD = 4.55$), respectively. Mean years of education completed for the YG and OD groups were 15.67 ($SD = 1.74$) and 15.53 ($SD = 2.31$), respectively. Groups differed significantly for age, $F(1, 85) = 4770.67, p < .001$, but not years of education, $F(1, 85) = .42, p = .75$. All participants completed the Mini-Mental Status Examination (Folstein, Folstein, & McHugh, 1975), a cognitive screening measure, and received scores of 26 or higher. Additionally, all participants passed hearing and vision screenings.

Language elicitation and transcription

Participants' discourse samples were collected in a single session. Samples consisted of the participants' description of two procedures, their description of the Nicholas and Brookshire (1993) single pictures, their storytelling of the wordless picture book Picnic (McCully, 1984), and recounts of three past experiences (see Table 1). Samples were digitally recorded and then orthographically transcribed in the Computerized Language Analysis software (CLAN) format (MacWhinney, 2000) by trained graduate assistants. Approximately 10% of the samples were randomly selected and transcribed again for reliability purposes. Intra- and inter-rater word-by-word transcription reliability was 96.12 and 95.2%, respectively. Nonwords and onomatopoeia were excluded via transcription codes in CLAN. Samples were further coded using the GEM command, which allows for marking and analyzing particular parts in samples.

Lexical Diversity

To investigate lexical diversity, D was calculated (Malvern & Richards, 1997) using the VOCD program in CLAN (MacWhinney, 2000) program. To calculate D , the program estimates the type/token ratio (TTR) for increasing number of tokens, starting with $N=35$ tokens to $N=50$ tokens (see McKee et al., 2000, for review) multiple times and estimates their average. Then using the least squares approach, a single curve is fitted to the empirically derived curve that is the function of a single value: D . See Figure 1 for an example of two speakers with different LD values as estimated by TTR and D .

Results

A 2 x 4 mixed analysis of variance was conducted to evaluate the effect of discourse type and age on lexical diversity (LD). The dependent variable was D . The between-participants factor was age group (young and old). The within-participants factor was type of discourse (procedures, single pictures, story, and recounts). The discourse type main effect and the discourse type by age interaction were tested using the multivariate criterion of *Wilk's lambda* (Λ). Significant results were found for the discourse type main effect, $\Lambda = .19, F(3, 82) = 118.90, p < .01, partial \eta^2 = .81$, and the interaction, $\Lambda = .90, F(3, 82) = 3.03, p = .03, partial \eta^2 = .10$. The age main effect was not significant, $F(1, 84) = 3.60, p = .06, partial \eta^2 = .04$.

To better understand the significant interaction, age and discourse type simple effects were examined by conducting a series of independent sample t-tests and paired-sample t-tests, respectively. To control for Type 1 error, alpha was set to .0125 and .025 for each age and discourse type simple main effect, respectively. Familywise error rate across the t-tests was controlled using the Holm's sequential Bonferroni approach.

Results indicated that the YG and OD groups demonstrated similar LD for the single pictures and story; however, the OD group yielded significantly greater LD for the procedures and recounts (see Figure 2). For both groups, procedures were associated with the least LD and recounts with the greatest. Single pictures and the story fell between, in that order. All comparisons were significant with the exception of the comparison of LD between single picture description and story telling for the YG group, $t(42) = 2.08, p = .04$.

Conclusions

The purpose of this study was to explore how discourse type influences lexical diversity and whether age-related differences exist. Four discourse types that are often used in clinical and research practice were compared using dedicated software, which allows for comparisons of LD regardless of length variation. From the results of the study, a LD hierarchy was found with the discourse tasks used; this hierarchy was similar for both older and younger participants. However, the older participants demonstrated significantly greater LD than the young participants for the procedures and recounts. This could be associated with the fact that procedures and recounts are not supported by pictorial stimuli.

The results of this study may carry methodological implications for discourse production sampling. Further, the findings add to our knowledge of how aging influences lexical diversity. Finally, it serves as the pivotal point for further exploring the nature of lexical diversity in healthy adults as well as adults with communication deficits by analyzing whether greater LD is associated with communicating pertinent information.

References

- Carroll, J. B. (1964). *Language and thought*. Englewood Cliffs, New Jersey.: Prentice-Hall.
- Engber, C. A. (1995). The relationship of lexical proficiency to the quality of ESL compositions. *Journal of Second Language Writing*, 4(2), 139-155.
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-mental state: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3), 189-198.
- Herdan, G. (1960). *Quantitative linguistics*. Butterworth, London:
- Jarvis, S. (2002). Short texts, best-fitting curves and new measures of lexical diversity. *Language Testing*, 19(1), 57-84.
- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk, vol 1: Transcription format and programs (3rd ed.)*. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- McCabe, A., & Bliss, L. S. (2003). *Patterns of narrative discourse*. Boston, MA: Pearson Education, Inc.
- McCully, E. A. (1984). *Picnic*. China: HarperCollins.
- McKee, G., Malvern, D., & Richards, B. (2000). Measuring vocabulary diversity using dedicated software. *Literary and Linguistic Computing*, 15(3), 323-337.
- Nicholas, L. E., & Brookshire, R. H. (1993). A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. *Journal of Speech & Hearing Research*, 36(2), 338-350.
- Phillips, L., & Jorgensen, M. W. (2002). *Discourse analysis as theory and method*. London: Sage Publication Ltd.
- Tweedie, F. J., & Baayen, R. H. (1998). How variable may a constant be? measures of lexical richness in perspective. *Computers and the Humanities*, 32(5), 323-352.
- Vermeer, A. (2000). Coming to grips with lexical richness in spontaneous speech data. *Language Testing*, 17(1), 65-83.
- Yule, G. U. (1944). *The statistical study of literary vocabulary*. New York, NY, US: Cambridge University Press.

Table 1

Tasks for each Discourse Genre

Discourse Genre	Tasks
Procedures	How to make a peanut butter and jelly sandwich How to plant a flower in a garden
Single Picture Description ^a	Cat on the Tree Birthday Cake
Story Telling ^b	Picnic
Recounts	Last Weekend Last Vacation Last Holiday

^aFrom Nicholas & Brookshire (1993) ^bMcCully (1984)

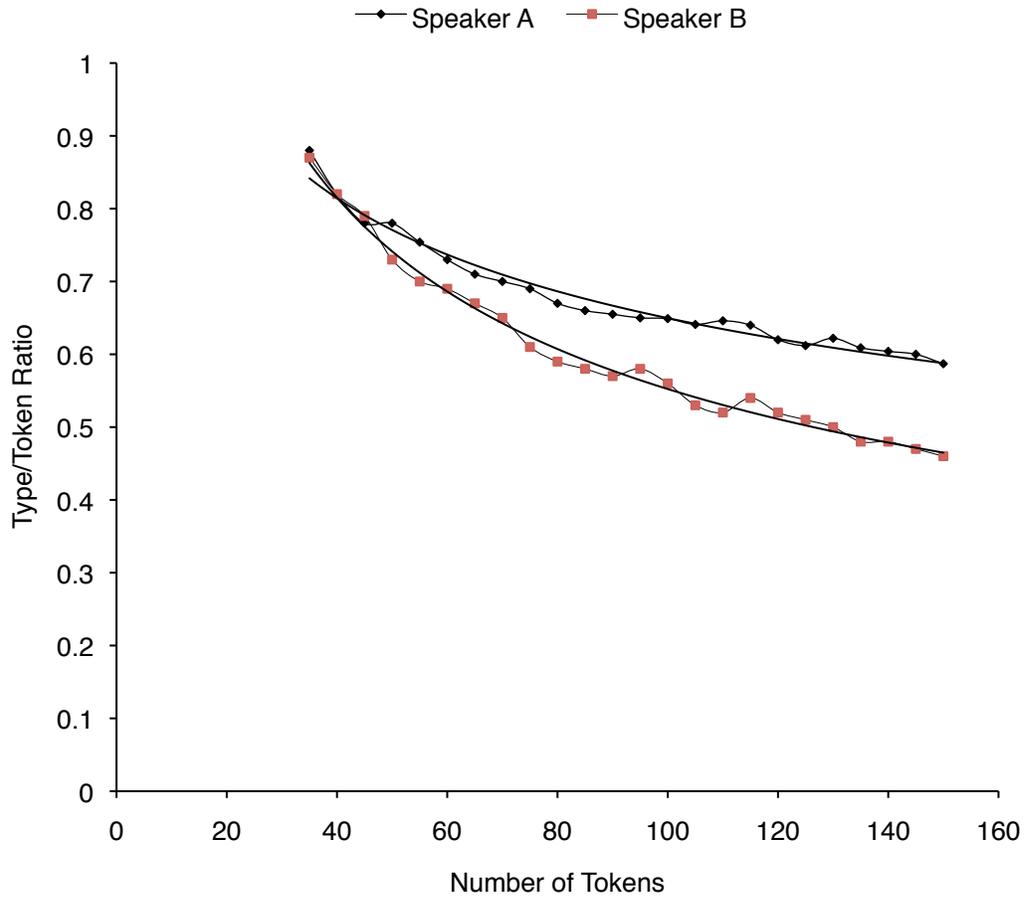


Figure 1. Average TTR values for increasing number of tokens for two speakers. The slope of the fitted line corresponds to different D values. The steeper the slope of the fitted line the lower the D value.

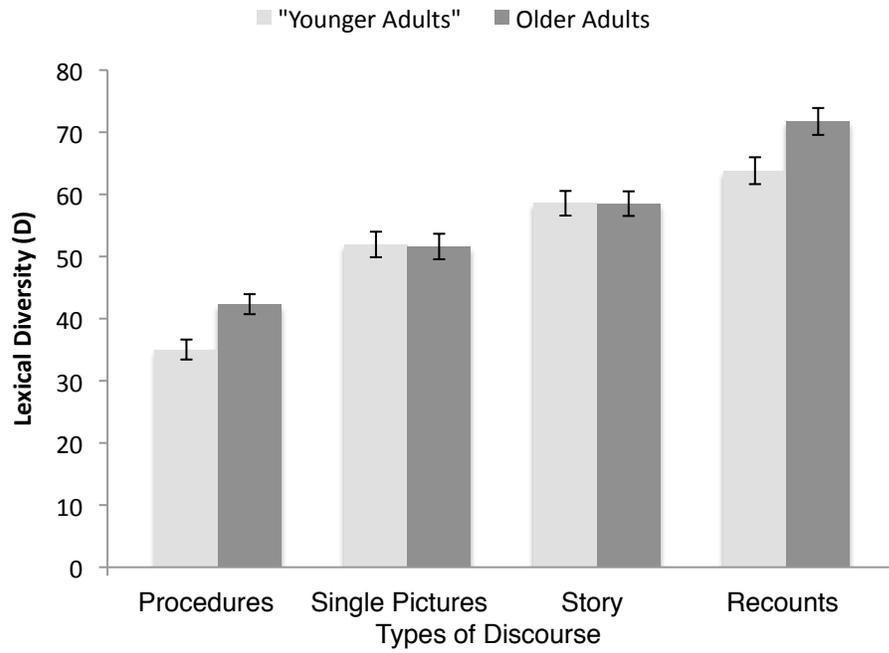


Figure 2. Mean lexical diversity (+SE) for the two age groups in four different discourse types. All differences were significant except for the comparison of LD between single picture description and story telling for the YG group.