Brain & Language 169 (2017) 57–63

Contents lists available at ScienceDirect

Brain & Language

journal homepage: www.elsevier.com/locate/b&l

Short communication

Longitudinal changes in linguistic complexity among professional football players

Visar Berisha [a,*], Shuai Wang [b], Amy LaCross [b], Julie Liss [b], Pamela Garcia-Filion [c]

[a] Department of Speech and Hearing Science and School of Electrical Computer and Energy Engineering, Arizona State University, United States
[b] Department of Speech and Hearing Science, Arizona State University, United States
[c] University of Arizona, College of Medicine-Phoenix, Dept of Biomedical Informatics, United States

ARTICLE INFO

Article history:
Received 19 July 2016
Revised 12 February 2017
Accepted 25 February 2017
Available online 17 March 2017

Keywords:
Chronic traumatic encephalopathy
Linguistic complexity
Natural language processing
Type-to-token ratio
Lexical density

ABSTRACT

Reductions in spoken language complexity have been associated with the onset of various neurological disorders. The objective of this study is to analyze whether similar trends are found in professional football players who are at risk for chronic traumatic encephalopathy. We compare changes in linguistic complexity (as indexed by the type-to-token ratio and lexical density) measured from the interview transcripts of players in the National Football League (NFL) to those measured from interview transcripts of coaches and/or front-office NFL executives who have never played professional football. A multilevel mixed model analysis reveals that exposure to the high-impact sport (*vs* no exposure) was associated with an overall decline in language complexity scores over time. This trend persists even after controlling for age as a potential confound. The results set the stage for a prospective study to test the hypothesis that language complexity decline is a harbinger of chronic traumatic encephalopathy.

1. Introduction

Repeated mild traumatic brain injury (mTBI), as experienced by professional athletes in high-impact sports, has been associated with an increased risk of developing chronic traumatic encephalopathy (CTE) Saulle & Greenwald, 2012. CTE is a progressive degeneration of brain tissue characterized by abnormal buildup of tau protein in neurons and glial cells (McKee et al., 2009). A conclusive diagnosis of CTE can be made only through post-mortem evaluation of p-tau deposits in the brain since behavioral manifestations of the disease overlap with other neurodegenerative disorders, such as Alzheimer's disease (Lenihan & Jordan, 2015; McKee et al., 2009).

The clinical presentation of CTE is diffuse and variable, but commonly includes disruptions in mood or behavior, cognitive decline, and/or motor impairments (see Lenihan & Jordan, 2015, for a review). In a review of 51 confirmed cases of CTE in athletes, McKee et al. find that one-third of the individuals were symptomatic when they retired from their sport and one-half showed signs of CTE within 4 years of retirement (McKee et al., 2009). Early and mild manifestations of these symptoms may exist unrecognized for years, masked by compensatory strategies or attributed to other etiologies or general personality traits. Yet early identification of probable CTE is critical for the development and testing of neuroprotective interventions. Fortunately, these subtle deficits can be revealed when pressure is exerted on cognitive resources.

The production of language, in the form of conversation or spontaneous written narrative, is a form of pressure on cognitive resources. It requires identifying words to express an idea, arranging these words in an order allowed by the language, all before initiating the articulatory muscles to even produce the speech, or fine motor control to write. As a result, a number of studies have demonstrated reductions in linguistic complexity of spoken and written discourse in patient populations with cognitive impairment and dementia (Kempler, 1995; Ripich, Vertes, Whitehouse, Fulton, & Ekelman, 1991; Roark, Mitchell, Hosom, Hollingshead, & Kaye, 2011; Snowdon et al., 1996). Our study is an extension of this work to a pre-symptomatic population at risk for CTE: players in the National Football League (NFL). We have compiled an extensive corpus of over 10,000 interviews with 10 NFL players (P), 9 of whom were on active rosters in the NFL as of the start of the 2016 season, and 18 NFL front office executives and coaches (C) who have never played professional football. This provides a rich and unique opportunity to evaluate longitudinal changes in language use for both groups (players and coaches/front-office executives) and to explore whether lexical complexity measures, derived

* Corresponding author at: Department of Speech and Hearing Science and School of Electrical Computer and Energy Engineering, Arizona State University, Tempe, AZ 85284, United States.

E-mail address: visar@asu.edu (V. Berisha).

http://dx.doi.org/10.1016/j.bandl.2017.02.003
0093-934X/

automatically from interview transcripts and tracked longitudinally, exhibit any decline.

There are a number of measures to estimate language complexity in the literature. These have been used to track second-language learning (Lu, 2012), child language development (Richards, 1987), and neurological health (Andreasen & Pfohl, 1976; Berisha, Wang, LaCross, & Liss, 2015; Borovsky, Saygin, Bates, & Dronkers, 2007; Bucks, Singh, Cuerden, & Wilcock, 2000; Fergadiotis, Wright, & West, 2013; Goldstein, Levin, Goldman, Clark, & Altonen, 2001; Kempler, 1995; Ripich et al., 1991; Roark et al., 2011; Snowdon et al., 1996; Tucker & Hanlon, 1998). Because this is an observational study, we focus on two simple, but robust, lexical complexity measures that have been established in other studies as potential markers of cognitive decline (Berisha et al., 2015; Kemper, Thompson, & Marquis, 2001; Le, Lancashire, Hirst, & Jokel, 2011): (1) lexical density (LD) – the ratio of the number of content words (verbs, nouns, adjectives, adverbs) to the total number of words in the text (Lu, 2012); (2) type-to-token ratio (TTR) - a ratio of the number of unique words in a text to the total number of words in the text. Qualitatively, the LD captures the density of ideas in a text and the TTR is a proxy measure for an individual's working vocabulary. Snowdon et al. related idea density in writing with cognitive health later in life (Snowdon et al., 1996). Roark et al. showed that the Content Density (a measure very similar to the LD used here) was a strong predictor of mild cognitive impairment (Roark et al., 2011). Similarly, the TTR has been used extensively by researchers as an analysis measure in a variety of clinical populations, including to assess linguistic differences between people with dementia associated with Alzheimer's and healthy controls (Bucks et al., 2000), to measure lexical diversity in aphasics (Fergadiotis et al., 2013), to assess the extent of lesions on speech production (Borovsky et al., 2007), and as a measure of linguistic diversity in preschool age children with language impairments (Richards, 1987). In our study we analyze longitudinal change in TTR and LD separately and compared between groups (players and coaches/front-office executives) using a multilevel mixed model approach.

Table 1
The table summarizes all players (top) and non-player personnel (bottom) that are included in the present study based on inclusion criterion described in the methods session. For the non-player personnel, we provide their age (as of the date of the first transcript), and the number of years of secondary education (college + postgraduate). For the players, we additionally include the number of years in the NFL (as of the date of the first transcript), their position on the field, and the average number of sacks per game (if Quarterback).

ID	Age	Education	Years in NFL	Position	Sacks/ game
Players (P)					
P1	29.8	4	7.1	Quarterback	1.79
P2	26.1	3	3.7	Nose Tackle	
P3	28.6	4	5.3	Quarterback	1.64
P4	23.7	4	0.27	Quarterback	2.51
P5	24.8	4	1.7	Cornerback	
P6	26.1	3	4.3	Wide Receiver	
P7	23.9	4	0.4	Quarterback	1.95
P8	29.7	4	6.4	Quarterback	1.81
P9	23.8	3	1.3	Quarterback	2.05
P10	23.3	3	1.1	Quarterback	2.74
Avg (St. dev.)	26.0 (2.6)	3.6 (0.5)	3.2 (2.5)		2.1 (0.4)

ID	Age	Education
Coaches + Front-Office Executives (C)		
C1	55.1	4
C2	57.9	4
C3	33.6	4
C4	36.1	4
C5	31.3	3
C6	63.0	4
C7	60.9	4
C8	50.0	5
C9	47.7	6
C10	42.8	6
C11	36.4	4
C12	63.6	2
C13	47.3	4
C14	52.8	4
C15	45.8	4
C16	58.7	5
C17	44.2	2
C18	36.9	4
Avg (St. dev.)	48.0 (10.4)	4.1 (1.0)

2. Results

In Table 1 we provide a list of individuals included in the study and summary statistics: the age (as of the date of the first interview), the number of years of college education, and, for the players, their position on the field, the number of years in the NFL (as of the date of the first interview), and the average number of times sacked per game (for the quarterbacks). An independent-samples *t*-test was conducted to evaluate differences in college education levels and in age between the two groups. There was no significant difference in college education between the two groups; $t(26) = 1.30$, $p = 0.21$. However, the C group mean age was significantly higher than that of the P group; $t(26) = 6.53$, $p < 0.0001$. For the C group, there was a positive correlation for LD with age [$r = 0.104$, $p < 0.001$]. For the P group, there was a negative correlation for TTR with age [$r = -0.425$, $p < 0.001$] and with number of years played [$r = -0.310$, $p < 0.001$].

In Table 2, we show the results of the mixed effects model. The analysis indicate that, compared to the coach/executive group, player status was associated with an overall decline in the TTR and LD scores over time. For TTR, there was a statistically significant decline over time [beta (SE): −0.016 (±0.002) points; $p_{unadj} = 0.012$]. The effect size increased [−0.027 (±0.003) points] after ($p_{adj} < 0.001$) adjusting for a potential confound (age at the time of first transcript). For LD, the decrease [−0.004 (±0.002) points] over time was statistically significant after controlling for age ($p = 0.012$) but not prior to [0.002 (±0.001); $p = 0.166$].

In Fig. 1, we show the slopes of the unadjusted model for the two complexity parameters plotted by group. Note that these measures represent the change in the complexity parameters over time and not their absolute values. A positive value indicates an increase in language complexity over time, whereas a negative value indicates a decrease in complexity over time. It is clear from the plot that there is a distinct difference in the trend exhibited by the two groups. For the P group, 70% of the individuals exhibited either a decline in LD or in TTR, and 40% exhibited a decline in both. For the C group, 44% of the individuals exhibited a decline in either parameter with 28% exhibiting a decline in both. As these figures show, this effect is more pronounced when only considering large negative changes in the language parameters (e.g. slopes < −0.0001).

To visualize data from individual subjects, Fig. 2 shows the TTR and LD plotted over time for two individuals – a coach and a player from the same team during the same period. We also show the line of best fit with 95% confidence bounds for each figure. The line was fit by converting the date to the number of days passed since 1/1/2005. These plots serve as exemplars indicative of the larger trend exhibited by the data. As the figure shows, both measures for the player exhibit a subtle decline over time; for the coach, the same parameters are increasing over time. While there are clear longitudinal trends in the plots, there is also a great deal of local variability. This is to be expected. Extensive work in the

Table 2
Effect of exposure to high impact sports (player) on language complexity scores over time, compared to front-office personnel with no exposure.

		Unadjusted			Adjusted[b]		
		Beta	SE	p-value	Beta	SE	p-value
TTR	Player[a]	−0.016	0.002	0.012	−0.027	0.003	<0.001
	Intercept	0.291	0.002	–	0.326	0.007	–
LD	Player[a]	0.002	0.001	0.166	−0.004	0.002	0.012
	Intercept	0.445	0.002	–	0.466	0.005	–

[a] Main effect of exposure to high impact compared to no exposure among coaches/executives.
[b] Adjusted for subject age at first transcript.

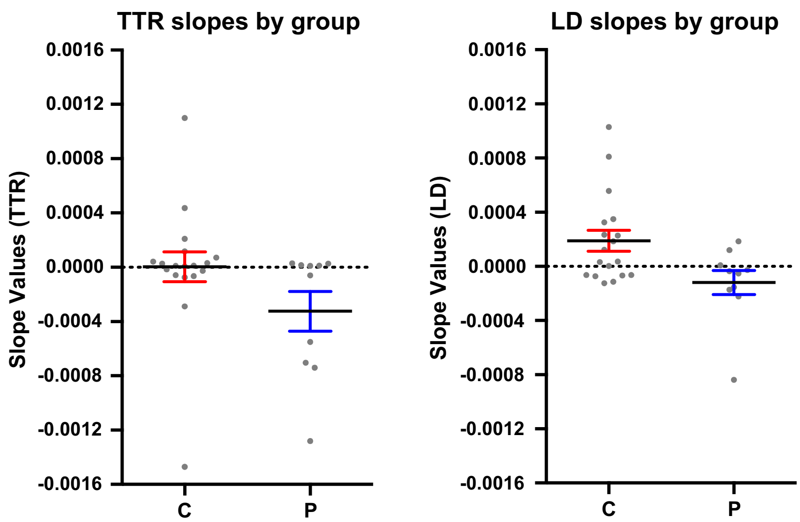

Fig. 1. A mixed effects model was used to analyze the between-group differences in the slopes of the TTR and LD parameters for the two groups. These figures show the individual slopes of the unadjusted model for the two language parameters (with standard error bars). Positive slope values indicate an increase in language complexity over time, whereas negative slope values indicate a decrease in language complexity over time.

variables influencing intra-speaker variation suggest that any level of language, be it phonological, phonetics, syntax, morphology, may be subject to both conscious/explicit and completely unconscious/subtle variation within a speaker, as conditioned by pragmatics, style-shifting, or register shifting (Coupland, 2007; Schilling, 2013).

In Fig. 3, we plot the LD slopes against the number of sacks/game [r = −0.740, p = 0.057] and the TTR slopes against the number of sacks/game [r = −0.630, p = 0.131] for the quarterbacks in the study. Neither result reached statistical significance, and additional data are required to determine whether the inverse trend between sacks/game and language decline holds.

3. Discussion

Despite the limitations inherent in a retrospective observational study, the finding of lexical complexity decline in these NFL players is remarkable. Numerous studies have shown that throughout adulthood, lexical complexity measures in neurologically healthy people remain stable or even increase with age (Kemper, 1990; Kemper, Greiner, Marquis, Prenovost, & Mitzner, 2001). For instance, it has been demonstrated that vocabulary increases over the lifespan (Wechsler, 2014). Kemper and Sumner have demonstrated that diversity and sophistication of lexical use increases with age (Kemper & Sumner, 2001). Another study identifies age-related increases in lexical complexity and vocabulary at the discourse level (Rabaglia & Salthouse, 2011). This means that, while the mean age of the P group was significantly lower than the C group at the time of the first interview, there is no reason to expect a decline in linguistic complexity for healthy younger controls.

It is only with advanced age or neurological disease or injury that we would expect decreases in the diversity and complexity of language when the cognitive-linguistic system is taxed (Berisha et al., 2015; Goldstein et al., 2001; Kemper, 1990; Kemper et al., 2001; Tucker & Hanlon, 1998). Reductions in the lexical complexity of spoken and written discourse have been demonstrated in numerous studies of cognitive impairment and dementia (Kempler, 1995; Ripich et al., 1991; Snowdon et al., 1996). This work dates back to Snowdon's et al. "*Nun Study*", which revealed that idea density and grammatical complexity in writings early in life were related to cognitive health later in life (Snowdon et al., 1996). Other studies have reported other forms of simplification, such as an over-reliance on the use of common phrases, non-specific indefinite nouns (e.g., *thing*), and high frequency/low-imageability verbs, and smaller vocabulary size (Berisha et al., 2015; Goldstein et al., 2001; Tucker & Hanlon, 1998). Against this backdrop, a finding of reduced lexical complexity in the P group warrants follow-up for cause.

There is reason to suspect that the changes we observe are due to exposure to the high-impact sport. The lexical complexity reductions for the majority of the NFL players sampled here is in keeping with results of other studies that have evaluated cognitive-linguistic function following mTBI. Relevant to the results of the current analyses, higher order linguistic processes such as lexical access, verbal fluency, verbal memory, as well as non-literal expression, and complex auditory comprehension have all been shown to be adversely affected by the occurrence of mTBI (Goldstein et al., 2001; Tucker & Hanlon, 1998; Vas, Chapman, & Cook, 2015). Additionally, it has been shown that discourse and narrative abilities (Borgaro, Prigatano, Kwasnica, & Rexer, 2003;

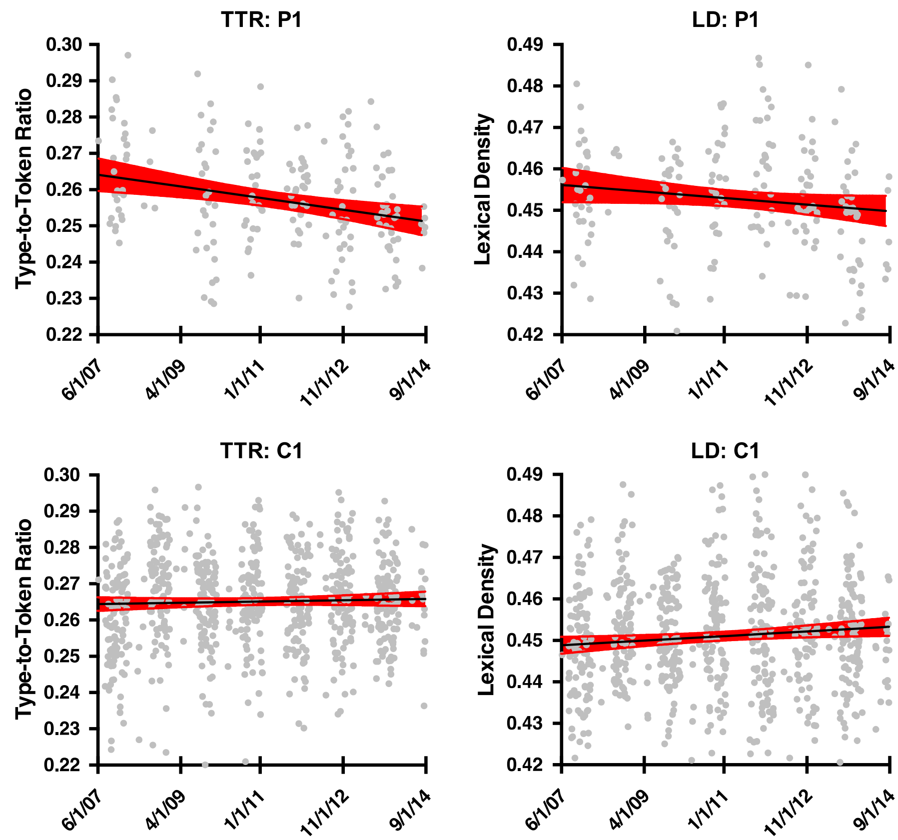

Fig. 2. The date-stamped average type-to-token ratio and lexical density plotted over time for C1 and P1 in the data. Since C1 and P1 are from the same team, we are able to plot their values over exactly the same time range.

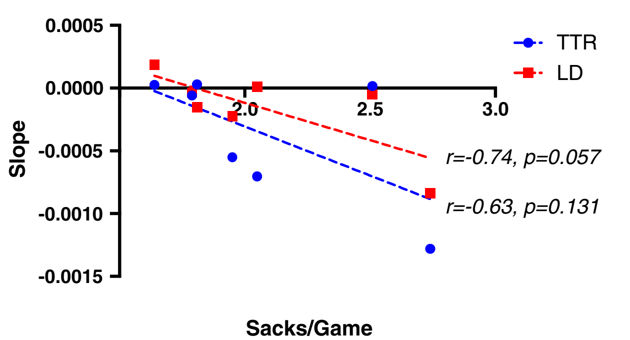

Fig. 3. The slope for each quarterback's lexical complexity parameter plotted against the average number of times sacked per game.

Marini et al., 2011), pragmatic ability (Douglas, 2010), and performance in naming tasks (King, Hough, Walker, Rastatter, & Holbert, 2006) decline in the event of mTBI. Supportive cognitive abilities such as memory are also adversely affected by mTBI and tasks requiring recall, organization, making inferences, and perception/discrimination all are shown to be affected (Borgaro et al., 2003; Levin, Eisenberg, & Benton, 1989). In addition, auxiliary analysis in our study reveals a relationship between the magnitude of change in TTR and LD and the number of hits a player experiences per game. To study this relationship, we focus on the 7 quarterbacks in our study. For this sub-group, we can approximate the frequency of hits they experience by estimating the average number of times per game that they are sacked by the opposing team. In Fig. 3, we show how the LD and TTR slopes vary with the number of sacks/game for the quarterbacks in the study. While this analysis does not yield statistically significant results, there seems to be an inverse trend between lexical complexity and the number of sacks per game: QBs with fewer sacks per game exhibit less pronounced decline. These negative correlations are at least partially driven by the QB with an especially high sacks/game measure and additional data from more subjects are required before a relationship between these variables is confirmed. However, taken together, these results and the existing knowledge base set the stage for a prospective study to test the hypothesis that language complexity decline is a harbinger of CTE.

A prospective study to test this hypothesis would allow for control over variables that may have influenced the present results. It would be preferable to compare P to an age-matched control group for longitudinal comparison in order to establish the healthy gradient for increased linguistic complexity over time in this population. Cognitive and functional impairments due to drug use (elicit abuse or prescriptions to manage known injuries or conditions) may be a possible confound in the subject group; as such, the study should attempt to account for these variables. Furthermore, such a study should match baseline language scores to ensure similar linguistic starting points across individuals. Ideally, we would establish baselines for an incoming rookie class and follow them longitudinally over many years. A prospective study in these populations should use question sets that are designed to elicit high complexity responses. Such questions would offer the additional advantage of placing greater cognitive-linguistic pressure on the participants to reveal smaller and/or earlier declines and would not be influenced by any prior media training. Furthermore, for all participants in the P group, the study should quantify the severity and frequency of hits in practice and during games. Finally, a careful power analysis should be conducted to determine the sample size prior to conducting the study. For the current study, a post hoc power analysis using a repeated measures study design indicated that the detected effect size for TTR had greater than 80% power whereas there was suboptimal power to detect a statistically significant effect size for the LD score. Our reported effect size for LD was much smaller than that for TTR; for small differences, a larger sample size is necessary to achieve statistical significance.

Without an in vivo marker of CTE, no study can conclude a causal relationship between the neural degeneration and clinical presentation during life. However, the data reveal trends for players that deserve causal explanation through prospective research designs. If it is found that declines in linguistic complexity are a marker of CTE, or track with other markers of CTE, this method offers a non-invasive and low cost solution. While the acquisition of longitudinal spontaneous speech samples is difficult retrospectively, prospective collection is a trivial matter with the proliferation of mobile devices and smartphones. Longitudinal behavioral studies aiming to identify early biomarkers of CTE or to distinguish between CTE and other tauopathies should also consider linguistic complexity as a sensitive component.

4. Methods

4.1. Data

Our study leveraged existing data, in the form of press conference transcripts, to quantitatively assess longitudinal changes in linguistic complexity in active NFL players. A subset of NFL teams makes publicly available interview transcripts from players and other team personnel (see the list teams in Appendix A). We collected these transcripts from 10 players and 18 coaches or front-office executives who have never played in the NFL.

Two aspects of these press conference transcripts are critical for the present analysis. First, the interviews are available over an extended period, allowing us to detect subtle within-person longitudinal changes in lexical complexity measures should they occur. Second, the interview format is non-scripted and spontaneous. This places the interviewee in the challenging position of needing to generate verbal responses that are appropriate and timely. With cognitive resources taxed, we would expect to see lexical complexity reductions to meet the interview timing demands, in the context of pathology or other mitigating circumstances (Berisha et al., 2015). Thus, the press conference transcripts provided the opportunity to track the longitudinal lexical complexity of NFL players at risk for CTE. We compare these changes against other individuals in the same setting but who have never played professional football, namely a group of NFL coaches and executives, for whom we were able to collect similar data.

We sampled a database of over 10,000 transcripts from press conference interviews, dating from 2007 to 2015. Each interview is date-stamped and consists of a spontaneous interaction between members of the press and each player (P) or executive or coach (C). The nature of the interviews is similar since the vast majority (>90%) come from either pre-game or post-game interviews. Thus, there is a great deal of overlap in the types of questions asked by the media. For the analysis, we focused only on the responses to the interviewer questions and generated a date-stamped concatenated longitudinal database of responses for each individual P or C. The details of the processing can be found in the data description in Appendix A.

4.2. Lexical complexity measures

From this merged data source, we analyzed two lexical complexity measures that have been shown to robustly change with cognitive decline (Berisha et al., 2015; Kemper, Thompson, & Marquis, 2001; Le, Lancashire, Hirst, & Jokel, 2011): (1) the type-to-token ratio (TTR) - a ratio of the number of unique words in a text to the total number of words in the text; (2) lexical density (LD) – the ratio of the number of content words (verbs, nouns, adjectives, adverbs) to the total number of words in the text (Lu, 2012). The lexical complexity measures are dependent on the length of the analyzed transcripts (Le et al., 2011); as a result, we adopted a fixed window length of 1000 words. From the concatenated transcripts of each person, we swept this 1000-word analysis window through the entire transcript and extracted the TTR and the LD (Bucks et al., 2000). This yielded a date-stamped list of lexical complexity measures. Prior to analysis, variability in the data was reduced by averaging LD and TTR parameters with the same date stamp.

4.3. Other variables

For every individual in the study, we used a combination of sources on the Internet to calculate their age. We did the same to find information regarding education levels. The number of years played in the NFL was calculated by comparing the date of the first transcript to the date they were drafted. The number of sacks per game was estimated as the total number of sacks[1] divided by the number of games played.

4.4. Statistical analysis

Linear regression analysis used a multilevel mixed model approach to examine the effect of exposure to high impact (player) [*versus* no exposure (coach/executives)] on the outcome TTR and LD scores. The model relied upon an autoregressive correlation structure and robust estimation of variance. For TTR and LD, separate models were constructed with the fixed effects of group (player *vs* front-office personnel), transcript number (time variable), and group-by-transcript number interaction and subject; and intercept as the random effect. The variable transcript number was centered on its mean. The working correlation structure for the within-subject repeated measurements was selected based on the lowest Akaike's Information Criteria. The regression estimates were adjusted for potential confounders: subject age at the time of the first transcript (all subjects). The slope estimates, interpreted as the average effect of high impact sports (players) on TTR and LD over time, are presented as the beta with standard error (±SE). The Pearson correlation coefficient between model slopes and the number of sacks/game was calculated for the quarterbacks in the study. For all analysis, statistical significance was defined as alpha of 0.05. Analyses were performed with STATA 13.1 (College Station, Tx).

Acknowledgements

The authors gratefully acknowledge partial support of this work by NIH-NIDCD R21 Research Grant 1R21DC013812, NSF STTR Phase I Grant 1448184 awarded to Aural Analytics, LLC, and an Arizona State University Athletics Research Grant.

Appendix A. Data source and pre-processing

A.1. Data source

The transcript data from all interviews were obtained from NFL team websites. A subset of NFL teams make publicly available websites that post transcripts from press conferences conducted by

[1] Number of sacks obtained from http://www.pro-football-reference.com.

players and other team personnel. These are typically under the headings of "Transcripts" or "News Conferences" on the website. Although the websites of NFL teams are similarly laid out, not all teams provide Question-Answer style transcripts for download. For this analysis, we focus on data from 11 NFL teams for which we could locate transcripts from interviews, each in a similar format. The table below lists the teams used in this study.

Team
New England Patriots
New York Giants
Washington Redskins
New York Jets
Atlanta Falcons
Seattle Seahawks
Baltimore Ravens
Cincinnati Bengals
Green Bay Packers
Houston Texans
Pittsburg Steelers

All data acquisition and natural language processing is done in Python. We downloaded all transcript data from the websites for each of the 11 teams, including the name of the interviewee, the date of the interview, and all questions and answers. The Hypertext Markup Language (HTML) in webpages was processed to extract the information corresponding to the interview. The resulting information was stored in text files for later processing. The output of this process is a series of text files, each containing the transcript of an interview on a date with an individual. The name of the file serves as an identifier for the individual and the date. The downloaded files were manually inspected by three volunteers to ensure that the content of the articles matched the identifier in the file name.

A.2. Data pre-processing

The first step in data pre-processing is removing extraneous information from each downloaded transcript. The original downloaded data contains both the questions asked and the answers provided by the interviewee. Occasionally, there are also opening statements from the interviewee before the Q&A session begins. We automatically pre-process the data to remove all questions and any opening statements so that only answers after questions are included in the processed text data for analysis. This was done to ensure that we capture only responses formed in real-time rather than prepared statements the individuals may have formulated prior to a press conference. Additionally, annotations in the transcripts (e.g. laughter) are removed from the analysis to focus only on the words of the interviewee. After pre-processing each transcript, what remains is a date-stamped stream of text spoken by the interviewee in response to specific questions asked by the media.

This initial processing results in a large set of transcripts for everyone in our database. The transcripts are of varying lengths; however, for each person in our database, we concatenate his transcripts into a single text stream, while preserving the date information for each portion of the stream that corresponds to a different transcript. We apply a sliding analysis window to the connected stream of texts for each person included in the analysis, selecting 1000 words at a time with a 30% overlap between two adjacent windows. A window of 1000 words has been used in other studies analyzing lexical complexity measures in dementia and Alzheimer's disease (Andreasen & Pfohl, 1976; Bucks et al., 2000). Each window may contain data from only a single transcript or can include data from multiple (shorter) transcripts. As a rule, we assign the date associated with the most recent transcript to each analysis window. The output of this step is a series of fixed-length, time-stamped windows from which we will extract the relevant lexical complexity measures.

We set two inclusion criteria for subjects in the study: The first criterion is that each player should have spoken at least 30,000 words. This ensures that we have at least 35 longitudinal samples for each individual for our analysis. The second criterion is that executives and coaches should have never played professional football during their career. After filtering based on these criteria, our database consists of data from 28 individuals: 10 players and 18 NFL coaches and executives personnel who have never played in NFL. Together they represent 11 out of the 32 NFL teams. After pre-processing, the language sample from each individual is greater than 2.1 h of continuous speech based on average speaking rate of 150 words per minute. Age and education (number of years in college) information for each individual is obtained from various online data sources, including the NFL team websites and the personal Wikipedia pages of each individual. The number of years played in the NFL was calculated by comparing the date of the first transcript to the date that they were drafted. The number of sacks per game was estimated as the total number of sacks (obtained from http://www.pro-football-reference.com) divided by the total number of games played.

References

Andreasen, N. J. C., & Pfohl, B. (1976). Linguistic analysis of speech in affective disorders. *Archives of General Psychiatry, 33*(11), 1361–1367.

Berisha, V., Wang, S., LaCross, A., & Liss, J. (2015). Tracking discourse complexity preceding Alzheimer's disease diagnosis: A case study comparing the press conferences of presidents ronald reagan and george herbert walker bush. *Journal of Alzheimer's disease: JAD, 45*(3), 959–963.

Borgaro, S. R., Prigatano, G. P., Kwasnica, C., & Rexer, J. L. (2003). Cognitive and affective sequelae in complicated and uncomplicated mild traumatic brain injury. *Brain Injury, 17*(3), 189–198.

Borovsky, A., Saygin, A. P., Bates, E., & Dronkers, N. (2007). Lesion correlates of conversational speech production deficits. *Neuropsychologia, 45*(11), 2525–2533.

Bucks, R. S., Singh, S., Cuerden, J. M., & Wilcock, G. K. (2000). Analysis of spontaneous, conversational speech in dementia of Alzheimer type: Evaluation of an objective technique for analysing lexical performance. *Aphasiology, 14*(1), 71–91.

Coupland, N. (2007). *Style: Language variation and identity*. Cambridge University Press.

Douglas, J. M. (2010). Relation of executive functioning to pragmatic outcome following severe traumatic brain injury. *Journal of Speech, Language, and Hearing Research, 53*(2), 365–382.

Fergadiotis, G., Wright, H. H., & West, T. M. (2013). Measuring lexical diversity in narrative discourse of people with aphasia. *American Journal of Speech-Language Pathology, 22*(2), S397–S408.

Goldstein, F. C., Levin, H. S., Goldman, W. P., Clark, A. N., & Altonen, T. K. (2001). Cognitive and neurobehavioral functioning after mild versus moderate traumatic brain injury in older adults. *Journal of the International Neuropsychological Society, 7*(3), 373–383.

Kemper, S. (1990). Adults' diaries: Changes made to written narratives across the life span. *Discourse Processes, 13*(2), 207–223.

Kemper, S., Greiner, L. H., Marquis, J. G., Prenovost, K., & Mitzner, T. L. (2001). Language decline across the life span: Findings from the Nun Study. *Psychology and Aging, 16*(2), 227.

Kemper, S., & Sumner, A. (2001). The structure of verbal abilities in young and older adults. *Psychology and Aging, 16*(2), 312.

Kemper, S., Thompson, M., & Marquis, J. (2001). Longitudinal change in language production: Effects of aging and dementia on grammatical complexity and propositional content. *Psychology and Aging, 16*(4), 600.

Kempler, D. (1995). Language changes in dementia of the Alzheimer type. *Dementia and Communication: Research and Clinical Implications*, 98–114.

King, K. A., Hough, M. S., Walker, M. M., Rastatter, M., & Holbert, D. (2006). Mild traumatic brain injury: Effects on naming in word retrieval and discourse. *Brain Injury, 20*(7), 725–732.

Le, X., Lancashire, I., Hirst, G., & Jokel, R. (2011). Longitudinal detection of dementia through lexical and syntactic changes in writing: A case study of three British novelists. *Literary and Linguistic Computing, 26*(4), 435–461.

Lenihan, M. W., & Jordan, B. D. (2015). The clinical presentation of chronic traumatic encephalopathy. *Current Neurology and Neuroscience Reports, 15*(5), 1–8.

Levin, H. S., Eisenberg, H. M., & Benton, A. L. (1989). *Mild head injury*. Oxford University Press.
Lu, X. (2012). The relationship of lexical richness to the quality of ESL learners' oral narratives. *The Modern Language Journal, 96*(2), 190–208.
Marini, A., Galetto, V., Zampieri, E., Vorano, L., Zettin, M., & Carlomagno, S. (2011). Narrative language in traumatic brain injury. *Neuropsychologia, 49*(10), 2904–2910.
McKee, A. C., Cantu, R. C., Nowinski, C. J., Hedley-Whyte, E. T., Gavett, B. E., Budson, A. E., ... Stern, R. A. (2009). Chronic traumatic encephalopathy in athletes: progressive tauopathy after repetitive head injury. *Journal of Neuropathology & Experimental Neurology, 68*(7), 709–735.
Rabaglia, C. D., & Salthouse, T. A. (2011). Natural and constrained language production as a function of age and cognitive abilities. *Language and Cognitive Processes, 26*(10), 1505–1531.
Richards, B. (1987). Type/Token Ratios: What do they really tell us? *Journal of Child Language, 14*(2), 201–209.
Ripich, D. N., Vertes, D., Whitehouse, P., Fulton, S., & Ekelman, B. (1991). Turn-taking and speech act patterns in the discourse of senile dementia of the Alzheimer's type patients. *Brain and Language, 40*(3), 330–343.
Roark, B., Mitchell, M., Hosom, J. P., Hollingshead, K., & Kaye, J. (2011). Spoken language derived measures for detecting mild cognitive impairment. *IEEE Transactions on Audio, Speech, and Language Processing, 19*(7), 2081–2090.
Saulle, M., & Greenwald, B. D. (2012). Chronic traumatic encephalopathy: A review. *Rehabilitation Research and Practice, 2012*.
Schilling, N. (2013). Investigating stylistic variation. *The handbook of language variation and change* (2nd ed., pp. 325–349).
Snowdon, D. A., Kemper, S. J., Mortimer, J. A., Greiner, L. H., Wekstein, D. R., & Markesbery, W. R. (1996). Linguistic ability in early life and cognitive function and Alzheimer's disease in late life: Findings from the Nun Study. *Jama, 275*(7), 528–532.
Tucker, F. M., & Hanlon, R. E. (1998). Effects of mild traumatic brain injury on narrative discourse production. *Brain Injury, 12*(9), 783–792.
Vas, A. K., Chapman, S. B., Cook, L. G. (2015). Language impairments in traumatic brain injury: A window into complex cognitive performance. *Handbook of clinical neurology* (pp. 128).
Wechsler, D. (2014). Wechsler Adult Intelligence Scale-Fourth Edition (WAIS–IV).