ARL-TN-0978 ● Nov 2019

Using Dragon for Speech-to-Text Transcription in Support of Human-Autonomy Teaming Research

by Andrea Krausman, Troy Kelley, Sean McGhee,
Kristin E Schaefer, and Sean Fitzhugh

NOTICES

Disclaimers

ARL-TN-0978 ● Nov 2019

Using Dragon for Speech-to-Text Transcription in Support of Human-Autonomy Teaming Research

Andrea Krausman, Troy Kelley, Kristin E Schaefer, and Sean Fitzhugh
Human Research and Engineering Directorate, CCDC Army Research Laboratory

Sean McGhee
SOS International LLC

REPORT DOCUMENTATION PAGE			*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)* November 2019	**2. REPORT TYPE** Technical Note	**3. DATES COVERED (From - To)** May 2019–September 2019
4. TITLE AND SUBTITLE Using Dragon for Speech-to-Text Transcription in Support of Human-Autonomy Teaming Research		**5a. CONTRACT NUMBER** **5b. GRANT NUMBER** **5c. PROGRAM ELEMENT NUMBER**
6. AUTHOR(S) Andrea Krausman, Troy Kelley, Sean McGhee, Kristin E Schaefer, and Sean Fitzhugh		**5d. PROJECT NUMBER** **5e. TASK NUMBER** **5f. WORK UNIT NUMBER**
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) CCDC Army Research Laboratory ATTN: FCDD-RLH-FC Aberdeen Proving Ground, MD 21005		**8. PERFORMING ORGANIZATION REPORT NUMBER** ARL-TN-0978
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		**10. SPONSOR/MONITOR'S ACRONYM(S)** **11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
ORCID ID: Kristin E Schaefer, 0000-0002-1342-3446

14. ABSTRACT
Dragon speech-to-text is an off-the-shelf transcription software developed by Nuance. Dragon version 15 was used to transcribe four prerecorded audio files of voice communication between two team members during a laboratory study to determine if this software is a feasible alternative to manual transcription to support human-autonomy teaming (HAT) research. In the present work, Dragon was able to transcribe the four audio files with 90% or better accuracy, without prior subject-voice training and using “out of the box” settings. Errors consisted of both incorrect word translations (e.g., “daring” translated as “during”, “teal” translated as “steel” or “chill”) and words that were missed by Dragon (i.e., words spoken but not picked up by Dragon). Missed words were most often conjunctions or connecting words (e.g., be). However, these errors did not seem to change the overall content of the communication. Transcription time with Dragon was, as expected, significantly reduced compared to manual transcription. For the longest audio file (15 min in duration) Dragon transcribed the file in 1 min, whereas manual transcription took 32 min. Overall, these results suggest that Dragon provides an efficient method for transcribing audio files and is a suitable platform to support HAT research. Efforts are currently underway to conduct real-time transcription with Dragon, thereby saving additional time and effort in the transcription process.

15. SUBJECT TERMS
verbal communication, team communication, speech-to-text transcription, Dragon software

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Andrea Krausman
Unclassified	Unclassified	Unclassified	UU	20	**19b. TELEPHONE NUMBER (Include area code)** (410) 278-5933

Contents

List of Figures

List of Tables

Summary

A goal of the US Army Futures Command and the CCDC Army Research Laboratory's Human-Autonomy Teaming Essential Research Program (HAT ERP) is to address the design and integration issues that arise when utilizing human-autonomy teams so that they can function effectively in complex environments. To accomplish this goal, we need to understand how humans and autonomy interact with one another and how critical team states such as team trust and cohesion are built and maintained in the context of human-autonomy teams. Of particular concern to this work is assessing team trust and cohesion through analyzing team member communication. There are a number of analytical methods developed for this purpose; however, one drawback is that they require audio data to be transcribed into text format prior to analysis. When transcription is performed manually, it is time consuming and costly, and as a result, researchers often shy away from performing in-depth analysis of communication data. Technological solutions, such as speech recognition or speech-to-text transcription, have gained traction in recent years, and although they significantly reduce the amount of time to transcribe audio data, accuracy is often not guaranteed. In the present work, Dragon, a commercial off-the-shelf transcription software developed by Nuance, was evaluated as a candidate for transcribing communication data to support the HAT ERP research and leverage communication data to analyze and extract information pertaining to team cohesion and trust.

Audio files from four team members, two male and two female, were transcribed both manually and using Dragon speech recognition software. Each audio file contained scripted speech and free conversation from a single team member. Word-for-word analysis showed that Dragon transcription times ranged from 38 s to 1 min, with accuracy of 90% or better, whereas manual transcription time ranged from 24 to 32 min, which is not surprising when considering that manual transcription requires multiple starts, stops, and rewinds to ensure accuracy. Audio files for this study contained one speaker, so it is likely that transcription time would be considerably longer if multiple speakers were present. Errors were primarily confusion over function words (e.g., the, an, as) or words translated into words that sound similar, so the main meaning or sentiment of the communication remained intact. These results suggest that Dragon provides a suitable alternative to manual transcription and can adequately support HAT research efforts to quantify team trust and cohesion. Future work is currently underway to use Dragon to support real-time transcription, thereby saving additional time and effort.

1. Introduction

In the future, the US Army proposes to utilize autonomous systems, such as the autonomy-enabled assets that comprise the Next Generation Combat Vehicle concept. The goal of which is to augment Soldier and team performance in increasingly complex environments. In theory, these types of autonomous systems can provide several advantages, such as reducing threats to Soldiers, providing additional "eyes" and "ears" to increase situation awareness, and even enable faster, more informed decision making. However, one challenge is to design and integrate the human-autonomy team in such a way that they are able to function effectively. Subsequently, there are several key research areas to be considered: 1) ensuring effective bi-directional communication that provides adequate situation awareness, 2) allowing human and autonomous team members to understand each other's intent, actions, and reasoning, and 3) engendering and maintaining critical team states, such as trust and cohesion.

One mechanism that enables us to better understand the underlying processes responsible for how a team functions is through their communication. Team communication is a rich source of information that can easily be obtained both in laboratory and field settings and through different analysis methods, and can shed light on the cognitions, emotions, and emerging team processes and states. For the purpose of this report, the focus is on verbal communication for crew coordination.

While technological advances make it easier to capture and record communication data in the laboratory and field settings, the challenge is transcribing the data into text form to support appropriate analysis techniques. This is an important consideration for the present human-autonomy team research since it will generate large volumes of rich audio communication data by multiple collaborative team members at one time that will broaden our understanding of how communication helps engender critical aspects of teamwork such as team trust, cohesion, and situation awareness. Essentially, transcription of audio data is done manually or by using transcription software. One major drawback to manual transcription is that it requires a significant investment of time and effort—in some cases it can take approximately 4 h for each hour of audio for an individual speaker (Britten 1995; Patton 2002) and can reach as high as 30 min of effort for 1 min of communication (Tiferes et al. 2016). Although software can transcribe in a fraction of the time, it is often untested for accuracy of complex communication. Recognizing the value of transcribing audio data and the drawbacks to manual methods, several attempts have been made to develop speech recognition systems and transcription software to assist in the process.

With respect to the present work, transcripts of communication between team members help identify the subject of communication (Tiferes et al. 2016), the type of statements being made (e.g., a question, instruction, comment, or answer; Burke et al. 2004), and quality of interaction, such as positive or negative reactions, acknowledgements, disagreements, and elaborations (Fischer et al. 2007; Rockmann and Northcraft 2010). Transcripts also provide the basis for extracting more in-depth information about team processes and emergent states, like trust and cohesion among team members. For example, qualitative coding, which involves a number of different approaches depending on the design and goals of the research (Saldaña 2013), can provide rich insights into team processes and emergent states by examining linguistic features such as word choice, sentence structure, and word counts. These can be analyzed to reveal similarities and differences between two or more pieces of transcribed text, thereby linking communication to cognitions, attitudes, and behaviors. Another important aspect of transcribed data is that it is easier to identify the source and destination of the communication. Subsequent analysis of these data can identify how information flows, or does not flow, both within and between teams, which can shed light on how strong or weak ties are between and within team members, as well as how these dynamics change over time (Baker et al. 2019).

Along with the qualitative techniques, there are methods that enable analysts to perform quantitative analysis on textual data (e.g., transcripts). One such method, the Linguistic Inquiry and Word Count (LIWC), is a computerized text analysis program that calculates linguistic patterns and the degree to which people use different categories of words that are linked to psychologically meaningful categories to assess factors such as attentional focus, emotionality, social relationships, thinking styles, and individual differences (Pennebaker et al. 2007). With respect to Human-Autonomy Teaming (HAT) research, LIWC could be used to assess the affective and social quality of team communication, which could then be linked to team cohesion or interpersonal trust.

There are also computational methods that can be used to extract meaning from team communication far more quickly than would be possible with human interpretation alone. Some of these approaches are outlined in Foltz and Martin (2008). For example, Latent Semantic Analysis (LSA) is a theory and method in computational linguistics that derives meaning from the context and pattern of word usage, irrespective of word meaning or syntax (for reviews, see Landauer et al. 1998; Dong 2005). LSA focuses on word co-occurrence and assumes that words that occur more frequently together are conceptually linked. When considering team cohesion, LSA can help shed light on the links between communication content and teamwork and has been useful in distinguishing high-performing and low-performing teams (Gorman et al. 2003; 2013).

Lastly, Social Network Analysis (SNA) involves the measurement and analysis of relational structures to provide insight into team communication, coordination, and consequently, performance. SNA is used to evaluate and visualize information transmission or communication throughout a network. This process results in a graphical representation of the qualitative aspects of communication (Pokorny et al. 2018), as well as links between network heterogeneity (i.e., team demographic diversity), network density (i.e., frequency of team communication), and team productivity (Reagans and Zuckerman 2001), or even organizational structure for mission operations (Fitzhugh and DeCostanza 2018).

One method of transcribing data into the proper format to support these analysis techniques is to use speech-to-text or speech recognition software. How we did this for this work is described in the next section. The remainder of this report documents the process and results of transcriptions of audio files using Dragon Naturally Speaking software, followed by a description of future work to examine real-time transcription.

2. Dragon Naturally Speaking

Dragon Naturally Speaking, a commercial off-the-shelf product, was evaluated as a candidate software to provide timely, accurate transcription of audio communication data recorded during human-autonomy teaming research. This software was selected because initial user testing of Dragon showed that it was a clear improvement over the Microsoft Speech-to-Text engines included in the Microsoft Windows Operating System. Additionally, it allows for a client-based system, which is separate from the Internet. This is an important consideration since the nature of HAT research often requires that all development work can be accomplished off-line in a secure environment. Additionally, given that communication data can be considered personally identifiable information, and in some military testing venues, all data, including communication data, could be considered classified information, a self-contained system, not connected to the web, to make translation decisions (like Google Cloud) was critical.

Further, although there are many transcription services available, the sheer volume of data generated by the 5-year HAT Essential Research Program (ERP) makes the cost of hiring a transcription service prohibitive. Among the other available transcription software options, Dragon is considered to be easy-to-use, cost effective (retail price is around $300), and it provides a real-time transcription option, which may help streamline the transcription process even further in the future, and a Batch Mode to reduce the time it takes to transcribe multiple audio files. Finally, Dragon version 15, used in this evaluation, is considered one of the best voice recognition software applications on the market (Allan and Turner 2019).

3. Methods

3.1 Audio Files

Audio files used for transcription were obtained from a distributed team communication study conducted at the CCDC Army Research Laboratory (Krausman 2019). During the study, dyads verbally shared and discussed pieces of information called factoids as they worked together to solve a fictitious terror plot. Factoids were presented to each team member in text format, similar to a script (Fig. 1). Team members verbally shared the scripted factoids with each another and engaged in a discussion (e.g., free conversation) as they worked together to solve the terror plot.

Four audio recordings of different lengths were used for this evaluation, two from male team members and two from female team members, to see if Dragon would respond differently to different voice types. Audio files were recorded as separate channels in .wav format using a Presonus 8-channel Mixer and Adobe Audition software. Prior to transcription, the audio files were exported into 16-bit PCM format using Audacity.

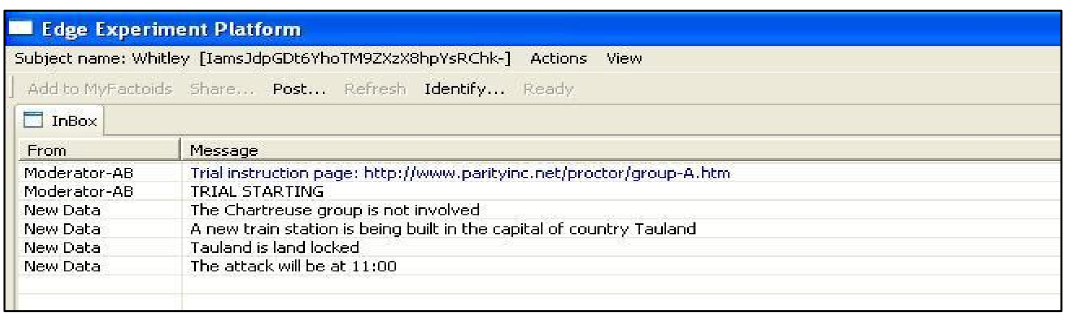

Fig. 1 Example of factoids verbally shared by team members

3.2 Manual Transcription Details

The audio recordings contained both scripted sharing of factoids, with the last few minutes devoted to free conversation. Both scripted speech and free conversation were transcribed manually and with the Dragon software. Manual transcription was accomplished by listening to the files and manually transcribing the content into text form. Accuracy was verified by listening to each audio file twice. Text from the manual transcriptions was imported into an Excel spreadsheet as individual words rather than the whole phrases shown in Fig. 1.

3.3 Dragon Transcription Details

Dragon Naturally Speaking version 15 was used "out of the box" for this evaluation to see how the software performed with minimal training. None of the settings provided in Dragon (e.g., languages, accents dialects) were required for this evaluation. Prior to transcription, Dragon was trained by using a generic user (e.g., a male voice), and this became the template or user profile for all subsequent transcriptions. The audio files used in this evaluation were recorded as part of a separate experiment, so Dragon was not trained on the voices of the individual team members. In addition, for the transcription, Dragon's pure dictation mode was used (which will ignore words it can interpret as commands, such as "copy that").

Each recorded audio file was transcribed separately by Dragon, and all output was saved as generic text files and imported into the same Excel spreadsheet as the manual output, as individual words rather than the whole phrases.

3.4 Word-for-Word Analysis

Prior to conducting the analysis, output from both transcriptions was arranged side-by-side in Excel columns. Next, words in both columns were formatted as lowercase and all punctuation marks were removed.* Using Excel functions (EXACT and IF), the words in the two columns were compared. If they were an exact match, they were scored as 1; if not, they were scored as 0. Finally, the percentage of correct matches was computed dividing the number of matches by the total number of words. Some of the factoid sets used in the experiment contained words that referred to fictitious locations (Omicronland, Piland, Spiderland, Perchland, etc.), so to ensure a fair analysis, these words were excluded from the analysis. Errors were defined as either Dragon reporting an incorrect word or failing to report a word or words that were spoken (i.e., a word spoken, but not appearing in the transcript).

4. Results and Conclusions

Descriptive statistics for the four transcriptions are shown in Table 1. Results showed that in all four cases, Dragon transcribed the audio conversations with 90% accuracy or better. As expected, Dragon was able to transcribe the files in a fraction of the time it took for manual transcriptions to be completed. Essentially, the

* Difference in punctuation placement can add false errors into the analysis.

manual transcription took two to three times longer than the total length of the audio files due to the tedious and time-consuming nature of manual transcription (playing audio, stopping, rewinding, playing again, etc.) and listening to the audio files twice to ensure the manual transcription was accurate. As mentioned previously, the time involved in manual transcription is a major drawback, especially when transcribing files with multiple speakers or low-quality audio (Britten 1995; Patton 2002).

Table 1 Descriptive statistics for Dragon transcriptions

Team member gender	Total words	Matches	Errors	Percent correct	File length (min)	Transcription time (min) (Manual)	Transcription time (min) (Dragon)
Female	431	402	29	93.3	14:46	32:28	1:02
Male	487	445	42	91.4	9:24	26:15	0:38
Female	391	360	31	92.1	10:45	24:47	0:58
Male	312	288	24	92.3	12:45	27:11	0:55

While the time it takes software or speech recognition technology to transcribe audio data is important, accuracy is also essential. For this evaluation, Dragon effectively transcribed audio data from both male and female team members, which suggests that even when using a default user profile, in this case a male voice for all transcriptions, and "out of the box" settings, Dragon still reached a high level of accuracy. As a result, it is unclear if training Dragon on individual team member voice profiles over time, which would use the pattern recognition features built-in to version 15 and used in this evaluation, would have resulted in even better accuracy. We investigate training effects further in the next section.

With respect to errors, the most common error Dragon made was confusing the words "and", "at", and "as". A second type of error was in words that sounded similar. For example, "teal" was transcribed as "steel", and "night" was transcribed as "light". However, there were a few instances where words were incorrectly transcribed (e.g., "operatives" was written as "offers" or "properties"); and where phrases were condensed into a single word (e.g., "and I'm only" was translated as "animal"). In addition to Dragon's transcription errors there were instances in which words were spoken but Dragon did not pick them up. For the most part, these "misses" were conjunctions (e.g., "and", "but", "if"), or smaller words connecting thoughts (e.g., "be"). When examining the errors Dragon made, it is important to consider that the audio files used for this evaluation were not specifically recorded for transcription purposes and, subsequently, likely contained noise on the channel, and the signal in several cases was quite low, which could be corrected for future analysis. It is also possible that pre-voice training would help resolve some of the instances of words being transcribed into words that sound similar to one another.

4.1 Effects of Training on Accuracy

To investigate the improvement in transcription accuracy with pre-voice training, we conducted a test of Dragon's accuracy with and without training. The training procedure entails reading a prompt provided by Dragon, and the training session ends when Dragon develops a profile for the speaker. In practice, this training took approximately 45 s. Because we could not use the audio files described in Section 3.1 for training, we evaluated accuracy on a test data set with a document containing excerpts from three corpora of English: iWeb, Wikipedia, and COCA (Corpus of Contemporary American English). These represent a variety of text types, including fiction, scientific reports, news articles, and transcripts of spoken language. Words in the corpus include a mix of ordinary language, technical terms (e.g., “computational phylogenetics”, “MacPherson strut”), proper nouns (e.g., “Harry Potter”, “Duke Ellington”, “Toyota Camry”), website names (e.g., “BreadExperience.com”, “Nuclear-Energy.net”), and acronyms (e.g., “NBA”). The entire corpus was 729 words and took approximately 3 min and 30 s to read aloud. For the nontraining condition, a male subject spoke directly into Dragon's Bluetooth headset, which relayed the audio directly to Dragon for live transcription. The training scenario followed the same approach but was preceded by using the training procedure.

To measure the accuracy of Dragon’s real-time transcription, we compared the transcripts produced under the training and nontraining cases to the original document. We tallied the number of correct words, modifications (mistranslations), deletions (spoken words that were not transcribed), and additions (words not spoken that were included in the transcription). We measured the accuracy of each transcription by calculating the word error rate (WER), which sums the modifications, deletions, and additions over the total number of words in the document. We report the results in Table 2. In the nontraining scenario, Dragon correctly translated 92.59% of words, a result consistent with our results in the team task; 72% of these errors were modifications, while 25% were deletions and 3% were insertions. In the training case, the accuracy rating rose to 94.65%, with modifications representing 82% of the errors and deletions accounting for the remaining 18%. Although the untrained Dragon transcription produced high accuracy, we found that the number of errors fell by 28% in the training scenario. The number of modifications fell by 18%, the number of deletions declined by 46%, and the number of additions dropped by 100%.

Table 2 Error rates with and without training

Scenario	Total words	Matches	Modifications	Deletions	Additions	WER
Nontraining	729	675	39	13	2	7.41%
Training	729	690	32	7	0	5.35%

Although we note improvement in the training case, the types of errors remained similar in both cases. Modification errors typically produced similar-sounding words, such as translating “and strut” to “in strip” in the nontraining condition and “instruct” in the training condition. Similarly, “poured in” was translated to “ported” in the nontraining condition and “part in” in the training condition. We also found a tendency for modification errors to occur when a prefix or suffix failed to translate. For example “relied” was transcribed as “lied”, and “cleared” was transcribed as “clear”. Notably, we did not find any modification errors on proper nouns or technical terms, although we did find a propensity for modification of website names: “HoopHall.com” was translated to “who Paul.com” in both cases and “BreadExperience.com” was translated to “read experience.com” in the nontraining scenario and “right experience.com” in the training scenario. Most of the deletion errors occurred when one-syllable words were not transcribed; this typically occurred with articles (e.g., “a” or “the”). Likewise, insertion errors, which only occurred twice, were single-syllable words: “is” and “work”. Although the nature of the errors was similar across training and nontraining cases, the volume of errors decreased notably during the training case.

4.2 Conclusions

Obviously, the best-case scenario would be 100% transcription accuracy when using technology, as this would make manual transcription a thing of the past. However, even though Dragon did not achieve 100% accuracy, the errors made did not appear to change the overall content and sentiment of the communication. Likewise, we found that training improved accuracy, but this improvement primarily occurred through proper translation of suffixes and prefixes as well as detection of articles. The content and sentiment of communication largely remained the same, but training helps to convey the linguistic structure more accurately. Therefore, it seems that the output from Dragon would be useful for further analysis and to help advance understanding of the link between communication and team trust and cohesion. Although perfect accuracy is likely in the future with advancements in artificial intelligence, deep learning, and processing speeds, for now, perhaps a combination of both methods is a reasonable compromise. For instance, having the text output from a Dragon transcription while listening to the file may result in gains in both transcription time and accuracy. In addition, there are a few recommendations that researchers and analysts should consider to ensure accuracy when using Dragon speech-to-text software to transcribe recorded audio:

- Use quality microphones.
- Boost the audio signal and reduce as much noise in the environment as possible.
- Where possible, record audio from each speaker on a separate channel, as transcription of multiple speakers becomes more difficult (Britten 1995; Patton 2002) whether it is done manually or using software.
- Utilize training to improve accuracy, given its minimal cost for subjects (less than a minute to complete). This is particularly useful for cases where omission of single-syllable words would be problematic, such as comparing usage of definite vs. indefinite articles.

In conclusion, Dragon version 15 provided timely and accurate transcription of the audio files used in the evaluation and will be a suitable platform for HAT ERP research examining the link between team communication and trust and cohesion. Additional work is currently underway to examine the effectiveness of the Dragon software for real-time transcription of live speech, with the intent to reduce transcription time while maintaining a high degree of accuracy.

5. References

Allan D, Turner B. Best voice recognition software of 2019. TechRadar; 2019 June 29 [accessed 2019 Nov 7]. https://www.techradar.com/best/best-voice-recognition-software.

Baker AL, Schaefer-Lay KE, Hill SG. Teamwork and communication methods and metrics for human-autonomy teaming. Aberdeen Proving Ground (MD): CCDC Army Research Laboratory (US); 2019 Oct. Report No. ARL-TR-8844.

Britten N. Qualitative interviews in medical research. British Medical Journal. 1995;311:251–253.

Burke JL, Murphy RR, Coovert MD, Riddle DL. Moonlight in Miami: a field study of human-robot interaction in the context of an urban search and rescue disaster response training exercise. Human Computer Interaction. 2004;19:85–116.

Dong A. The latent semantic approach to studying design team communication. Des Stud. 2005;26(5):445–461.

Fischer U, McDonnell L, Orasanu J. Linguistic correlates of team performance: toward a tool for monitoring team functioning during space missions. Aviation Space Environ Med. 2007;78(5):B86–95.

Fitzhugh SM, DeCostanza AH. Procure, persist, perish: communication tie dynamics in a disrupted task environment. Soc Netw Anal Min. 2018;8(1):37.

Foltz PW, Martin MJ. Automated communication analysis of teams. In: Salas E, Goodwin GF, Burke CS, editors. Team effectiveness in complex organizations. Cross-disciplinary perspectives and approaches. Abingdon (UK): Routledge; 2008. p. 411–431.

Gorman JC, Foltz PW, Kiekel PA, Martin MJ, Cooke NJ. Evaluation of latent semantic analysis-based measures of team communications content. Proc Hum Factors Ergon Soc. 2003;47(3):424–428.

Gorman JC, Martin MJ, Dunbar TA, Stevens RH, Galloway T. Analysis of semantic content and its relation to team neurophysiology during submarine crew training. Berlin (Germany): Springer; 2013.

Krausman AS. The impact of communication delays on distributed team interaction [unpublished dissertation]. [Blacksburg (VA)]: Virginia Tech; 2019.

Landauer TK, Foltz PW, Laham D. An introduction to latent semantic analysis. Discourse Process. 1998;25(2–3):259–284.

Patton MQ. Qualitative research and evaluation methods. 3rd ed. Thousand Oaks (CA): SAGE Publications; 2002.

Pennebaker JW, Booth RJ, Francis ME. Linguistic inquiry and word count; 2007. http://www.liwc.net/.

Pokorny JJ, Norman A, Zanesco AP, Bauer-Wu S, Sahdra BK, Saron CD. Network analysis for the visualization and analysis of qualitative data. Psychol Methods. 2018;23(1):169–183.

Reagans R, Zuckerman EW. Networks, diversity, and productivity: the social capital of corporate R&D teams. Organ Sci. 2001;12(4):502–517.

Rockmann KW, Northcraft GB. Expecting the worst? The dynamic role of competitive expectations in team member satisfaction and team performance. Small Group Res. 2010;41(3):308–29.

Saldaña J. The coding manual for qualitative researchers. 2nd ed. Los Angeles (CA): SAGE Publications; 2013.

Tiferes J, Hussein AA, Bisantz A, Kozlowski JD, Sharif MA, Winder NM, Allers AN, Cavuoto L, Guru KA. The loud surgeon behind the console: understanding team activities during robot-assisted surgery. J Surg Educ. 2016;73(3):504–512.

List of Symbols, Abbreviations, and Acronyms

ERP	Essential Research Program
HAT	Human-Autonomy Teaming
LIWC	Linguistic Inquiry and Word Count
LSA	Latent Semantic Analysis
PCM	Pulse-code Modulation
SNA	Social Network Analysis
WER	word error rate

1 (PDF) DEFENSE TECHNICAL
INFORMATION CTR
DTIC OCA

1 (PDF) CCDC ARL
FCDD RLD CL
TECH LIB

1 (PDF) CCDC ARL
FCDD RLH B
T DAVIS
BLDG 5400 RM C242
REDSTONE ARSENAL AL
35898-7290

1 (PDF) CCDC ARL
FCDD HSI
J THOMAS
6662 GUNNER CIRCLE
APG, MD 21005-5201

1 (PDF) USAF 711 HPW
711 HPW/RH K GEISS
2698 G ST BLDG 190
WRIGHT PATTERSON AFB OH
45433-7604

1 (PDF) USN ONR
ONR CODE 341 J TANGNEY
875 N RANDOLPH STREET
BLDG 87
ARLINGTON VA 22203-1986

1 (PDF) USA NSRDEC
RDNS D D TAMILIO
10 GENERAL GREENE AVE
NATICK MA 01760-2642

1 (PDF) OSD OUSD ATL
HPT&B B PETRO
4800 MARK CENTER DRIVE
SUITE 17E08
ALEXANDRIA VA 22350

13 (PDF) CCDC ARL
FCDD RLH
J LANE
Y CHEN
P FRANASZCZUK
K MCDOWELL
K OIE
K SCHAEFER
FCDD RLH BD
D HEADLEY
FCDD RLH FA
A DECOSTANZA
FCDD RLH FB
A EVANS
FCDD RLH FC
J GASTON
A KRAUSMAN
FCDD RLH
A MARATHE
FCDD RLH FD
S MCGHEE