

An Exploration of Off Topic Conversation

Whitney L. Cade
University of Memphis
365 Innovation Drive
Memphis, TN 38152-3115
wlcade@memphis.edu

Blair A. Lehman
University of Memphis
365 Innovation Drive
Memphis, TN 38152-3115
baleh-
man@memphis.edu

Andrew Olney
University of Memphis
365 Innovation Drive
Memphis, TN 38152-3115
aolney@memphis.edu

Abstract

In a corpus of expert tutoring dialogue, conversation that is considered to be “off topic” (non-pedagogical) according to a previous coding scheme is explored for its value in tutoring dynamics. Using the Linguistic Inquiry and Word Count (LIWC) tool, phases of tutoring categorized as “off topic” were compared with interactive problem solving phases to explore how the two differ on the emotional, psychological, and topical dimensions analyzed by LIWC. The results suggest that conversation classified as “off topic” serves as motivation and broad pedagogy in tutoring. These findings can be used to orient future research on “off topic” conversation, and help to make sense of both previous coding schemes and noisy data sets.

1 Introduction

Methods of investigating a large and noisy data set are of paramount importance in computational linguistics. Quite often, qualitative coding schemes are used to capture snapshots of the data set, but these may gloss over finer details or miss the larger picture. Add to that the messy and unpredictable nature of naturalistic data, and analysis becomes even more complicated. Therefore, a multi-method approach to understanding pre-existing coding schemes and orienting future in-depth analyses of those schemes proves to be a useful means of exploring one’s data.

Dialogue, particularly tutorial dialogue, is one area where large, noisy data sets are common. Computer and human tutoring data have been

parsed, coded, and tested by a number of researchers, and much effort has been put into making sense of the variability in the task-oriented dialogue (e.g. Chi, Roy, and Hausmann, 2008; Graesser, Person, and Magliano, 1995; Person, Lehman, and Ozburn, 2007). This work has all been in pursuit of a deep understanding of the complex interaction between the human tutor and student, which, if understood, could be used to boost the efficacy of artificially intelligent computer tutors. Expert human tutoring has been found to increase learning gains by as much as 2 sigmas (Bloom, 1984), which makes understanding their methods and motives the goal of any tutor research.

The corpus under examination here was collected with the express purpose of understanding how truly expert tutors manage a tutoring session, with an emphasis on creating a corpus of naturalistic dialogue data. The corpus has been investigated at two different grain sizes, a dialogue move level and a sustained phases level. Our study investigates in detail an “other” category that these coding schemes, which emphasize the pedagogy of the tutors and the students reactions, classify as “off topic” conversation. Off topic conversation, by virtue of its name, does not address the tutoring task in which the tutor and student are engaged. However, given the prevalence of off topic conversation in the corpus, it is perhaps more likely that the function or utility of off topic conversation in expert tutoring is indirect rather than non-existent, suggesting that the noisiest part of the tutoring dialogue corpus, off topic conversation, should be further explored.

Because any topic not pertaining to the topic at hand may be broached in off topic conversation and because the dialogue itself is full of false

starts, interruptions, and fragmented sentences, it is reasonable to explore off topic conversation using a bag of words method that is applicable to a variety of formal and informal texts. One such method is the Linguistic Inquiry and Word Count (LIWC) tool developed by Pennebaker et al., (2001), which looks for words that fall into specific, predetermined categories such as COGNITIVE MECHANISMS and POSITIVE EMOTIONS, then reports the percent of words in the document that fall into that category. LIWC provides over 70 possible categories, and can help sketch a rough picture of the verbal dynamics of a text (Mairesse and Walker, 2006; Mihalcea and Strapparava, 2009). Using a readily available tool like LIWC allows an examination of the variability within off topic conversation based on predetermined LIWC features. We can also compare these results to a prominent pedagogical category, such as scaffolding, that a current coding scheme particularly emphasizes, and examine the differences between the two.

In this analysis, the task-orientation and utility of “off topic” conversation are investigated by comparing its outcome scores in certain dimensions of LIWC to a classic pedagogical and interactive phase of tutoring: scaffolding (Rogoff and Gardner, 1984). Scaffolding, previously identified in a tutorial dialogue coding scheme (Cade, Copeland, Person, and D’Mello, 2008), involves much of the conversational give-and-take expected in casual off topic conversation, but is considered to be a very focused, on task phase of tutoring. Knowing how off topic conversation differs from scaffolding may help further exploration of this forgotten phase of tutoring. Likewise, it would give us direction in how to structure future coding schemes that would help bring clarity to the data set.

2 Methods

In this study, pedagogical and non-pedagogical phases of expert tutoring sessions were compared on linguistic dimensions to get at the diverse nature of off topic conversation within a naturalistic expert tutoring session.

The corpus under examination was collected in a previous study on expert human tutors. Therefore, what follows is a brief synopsis of how this corpus was collected.

Ten expert math and science tutors (4 male and 6 female) were recruited through local tutoring

agencies and schools. Tutors were considered “expert” when they met the following criteria: they had to be licensed to teach at the secondary level, have five or more years of tutoring experience, be employed by a professional tutoring agency, and come highly recommended by school personnel who specialize in providing support to students who are struggling academically. Student participants were in grades 7 to 12, except for one who was obtaining a GED. All of the students were in academic trouble and actively sought out tutoring.

All sessions were unobtrusively videotaped at the location decided upon by the tutor and student. The researcher turned on the camera and left the room when the session began. Each student participated in a maximum of two tutorial sessions, while each tutor participated in between two and eight tutoring sessions. These 50 1-hour tutoring sessions were then transcribed.

Two previously identified phases of tutoring (or “modes”), Off Topic and Scaffolding, were compared to investigate their psychological, emotional, and topical differences. To do this, instances of each mode were extracted from 30 sessions (all sessions that contained at least one Off Topic and one Scaffolding mode). If a session had multiple occurrences of a single mode, those modes were compiled into a single document. Documents were capped at 1000 words each to prevent differences in word count between the modes from affecting the outcomes. These documents were also separated by speaker (tutor or student); speakers may be differentially motivated to broach certain topics, and so separating out these effects leads to more specific identification of conversational dynamics. Each session’s Scaffolding and Off Topic document was then analyzed using LIWClite 7, which calculates the percentage of each document’s words that fall into specific, predefined categories. Though this version of LIWC offers over 70 linguistic categories, only 15 were of interest in determining the nature of off topic conversation: SOCIAL PROCESSES (ex: *mate, talk, they*), FAMILY (*daughter, husband, aunt*), FRIENDS (*buddy, neighbor*), AFFECTIVE PROCESSES (*happy, cried*), POSITIVE EMOTION (*nice, sweet*), NEGATIVE EMOTION (*hurt, ugly, nasty*) ANXIETY (*worried, nervous*), TENTATIVENESS (*maybe, perhaps*), CERTAINTY (*always, never*), WORK (*majors, class*), ACHIEVEMENT (*earn, win*), LEISURE (*chat, movie*), HOME (*kitchen, family*), NONFLUENCIES (*umm,*

hm), and FUTURE (*will, gonna*).

These categories are the most relevant in illustrating the emotional, topical, and psychological picture of conversation in tutoring when compared with the more on-task behavior of problem solving.

3 Discussion of Results

LIWC Category	T /S	Off Top M	Scaff M	Wilcoxon p-val	Paired t-test t-val	Cohen's d
Social Process	T	11.15	7.75	<0.01	<0.01	1.37
	S	8.25	4.87	<0.01	<0.01	0.90
Positive Emotion	T	5.41	4.83	0.27	0.29	
	S	6.54	4.54	0.09	0.05	0.47
Tentative	T	3.10	1.91	<0.01	<0.01	1.08
	S	2.68	1.60	0.02	0.02	0.65
Work	T	2.90	1.10	<0.01	<0.01	0.86
	S	2.70	2.09	0.54	0.43	
Achieve	T	1.02	0.95	0.67	0.76	
	S	0.52	1.89	<0.01	<0.01	-0.92
Leisure	T	0.78	0.23	0.60	0.27	
	S	0.50	0.15	0.05	0.07	0.50
Home	T	0.30	0.04	0.02	0.05	0.53
	S	0.24	0.01	0.03	0.17	0.37
Nonfluen.	T	1.51	1.11	0.04	0.08	0.44
	S	3.89	4.14	0.17	0.82	
Future	T	1.13	1.23	0.80	0.66	
	S	0.74	1.35	0.01	0.04	-0.49

Table 1. LIWC Dimensions with Significant Results

Since a normal distribution of scores cannot be assumed in this analysis, comparisons between Off Topic conversation and Scaffolding dialogue were made by comparing the LIWC scores of the modes using both Wilcoxon's signed-rank test and a paired t-test, with similar outcomes. Effect sizes were also analyzed by calculating Cohen's *d*. Table 1 illustrates the significant results that emerged. In total, each category investigated occurs more in Off Topic than in Scaffolding, with the exception of a student's discussion of ACHIEVEMENT and FUTURE.

From this analysis, an interesting pattern of results emerges. The Off Topic mode had previously been characterized as a conversation that had nothing to do with the lesson at hand, which connoted that it is fairly irrelevant. However, Off Topic does not seem to be so wholly "off topic." Tutors and students in the Off Topic mode talk about work more often than they do in the Scaffolding mode, which is a mode where nothing but work is done. WORK words, according to the authors of LIWC, are mostly school-related. Off Topic may be a mode that allows the tutor to discuss test-taking skills, study strategies, and remind students what

tasks need to be completed before the next tutoring session. For instance, one tutor divided up a study guide into manageable portions that needed to be completed every night so that the student would be prepared for an upcoming test. Previous to now, these conversations have only been qualitatively observed, but this supports a more in-depth analysis of what type of work tutors are talking about when they are supposedly discussing non-pedagogical topics.

This hypothesis is supported by the significant amount of conversation that takes place in Off Topic about the home; if FAMILY and FRIENDS (which may crop up in casual conversation about HOME-related topics) are not discussed significantly more in Off Topic, but HOME is, it may be that tutors are informing students of what sort of work needs to be done at home, and strategies to get work completed when on their own.

This may also explain why both students and tutors use more TENTATIVE words in Off Topic. Although it would seem that students should be more tentative and nonfluent when discussing difficult problem solving, they may be tentative in Off Topic when the tutor makes suggestions about studying and working. These suggestions of the tutor's may be framed using language like "maybe" and "perhaps" to make them more polite, and the student echoes this language in return. Thus, tentativeness may not come from uncertainty, but from suggestions couched in polite language.

It also appears that Off Topic conversation may not serve as a "pep talk" time; although it does contain more POSITIVE EMOTION words than Scaffolding, it does not expound upon the student's achievements. ACHIEVEMENT words are more common in Scaffolding, where students are receiving praise for their problem solving efforts. Off Topic conversation may seek to motivate the student in more subtle ways. By using more words that refer to SOCIAL PROCESSES (such as the third person plural and words like "talked"), the tutor and student may be building rapport with one another. This rapport may become important later on when the tutor gives the student blatantly negative feedback (Person et al., 2007), which can be motivationally damaging. Rapport may protect against flagging motivation in the student when the tutor uses "us" language and connects with the student in a more casual conversation.

4 Conclusions and Future Work

Our goal in this work was to use a simple linguistic analysis tool to uncover the hidden depths of an existing dialogue coding scheme. The use of such tools can paint a rich picture of the psychological, emotional, and topical content of a corpus, and can be used in two ways: first, it may help determine if a deeper inquiry into a hypothesis is warranted, and second, it can immediately orient future research towards key issues in a corpus without the less rigorous speculation and qualitative observations. The nature of broader coding schemes can come to be understood in a multifaceted manner using linguistic analysis, which may also inform future work.

Here, we have observed that off topic conversation in an expert tutoring dialogue corpus operates in a multidimensional way that is not irrelevant when studying the dynamics of an expert tutoring session. By using the LIWC tool developed by Pennebaker et al. (2001), themes concerning interpersonal rapport and global pedagogy emerge. The purpose of “off topic” conversation in tutoring may therefore be linked more to building a relationship between the tutor and the student, which is necessary for the trials of problem solving, and for the dispensation of “study strategies” that are more globally task-oriented, but are, nonetheless, important in understanding the pedagogical strategies of expert tutors. Off topic conversation was also hypothesized to function similarly in other tutorial work (Rosé, Kumar, Alevén, Robinson, and Wu, 2006).

One way of adding validity to these claims would be to investigate the topics broached in Off Topic through a topics model. In this way, recurring themes in off topic conversation can be revealed, and these themes can be aligned with the LIWC findings to see if a pattern emerges. From there, a new coding scheme may be devised to capture the multiple types of off topic conversation, which, for now, seem to be divided between interpersonal, rapport building and global pedagogy. This method of exploring a corpus has proven to be a useful approach when investigating possible avenues of improvement to coding schemes.

Acknowledgements

The research reported here was supported by the Institute of Education Sciences, U.S. Department

of Education, through Grant R305A080594 to the University of Memphis. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

References

- Benjamin Bloom. 1984. The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, 13:4-16.
- Whitney Cade, Jessica Copeland, Natalie Person, and Sidney D’Mello. 2008. Dialogue modes in expert tutoring. *Proceedings of the 9th International Conference on Intelligent Tutoring Systems*, 470-479. Springer-Verlag, Berlin, Germany.
- Micheline Chi, Marguerite Roy, and Robert Hausmann. 2008. Observing tutorial dialogues collaboratively: Insights about human tutoring effectiveness from vicarious learning. *Cognitive Science*, 32(2):301-341.
- Art Graesser, Natalie Person, and Joseph Magliano. 1995. Collaborative dialogue patterns in naturalistic one-on-one tutoring. *Applied Cognitive Psychology*, 9:359-387.
- François Mairesse and Marilyn Walker. 2006. Automatic Recognition of Personality in Conversation. In *Proceedings of the Human Language Technology Conference of the North American Chapter of the ACL*, 85–88. Association for Computational Linguistics, New York.
- Rada Mihalcea and Carlo Strapparava. 2009. The Lie Detector: Explorations in the Automatic Recognition of Deceptive Language. In *Proceedings of the ACL-IJCNLP 2009 Conference Short Papers*, 309-312. Association for Computational Linguistics, Suntec, Singapore.
- James Pennebaker, Martha Francis, and Roger Booth. 2001. *Linguistic Inquiry and Word Count (LIWC): LIWC2001*. Lawrence Erlbaum Associates, Mahwah, NJ.
- Natalie Person, Blair Lehman, and Rachel Ozbun. 2007. Pedagogical and motivational dialogue moves used by expert tutors. Presented at the 17th Annual Meeting of the Society for Text and Discourse. Glasgow, Scotland.
- Barbara Rogoff and William Gardner. 1984. Adult guidance of cognitive development. *Everyday cognition: Its development in social context*, 95-116. Harvard University Press, Cambridge, MA.
- Carolyn Rosé, Rohit Kumar, Vincent Alevén, Allen Robinson, & Chih Wu. 2006. CycleTalk: Data driven design of support for simulation based learning. *International Journal of Artificial Intelligence in Education*, 16:195-223.