ISBN: 978-989-8704-22-1 © 2020

ON-LINE COLLABORATIVE KNOWLEDGE BUILDING IN HIGHER EDUCATION: TESTING A QUANTITATIVE MODEL

Ian S. McGowan
University of Technology, Jamaica

ABSTRACT

Upto now, the knowledge building influence of the fundamental communicative functions during an on-line collaborative learning (OLCL) session, i.e. argumentative, responsive, elicitative, informative and imperative have been mainly based on results from qualitative studies, results that could have been strengthened by quantitative approaches. Starting from a literature review, we formulate a dual quantitative model of an on-line collaboration knowledge building (OCKB) that described these communicative functions, and aim to validate this model in a computer science topic related OCKB with a total of n=44 participants. Corpuses are collected for manual dialog act coding and communicative function variable calculations. A regression analysis failed to provide for the hypothesized model on seven of the eight tests on the basis of quantitative data. Findings suggest the imperative communication function best explain the assessment results statistically alone and in some significance in combination with some of the other communicative functions.

KEYWORDS

Knowledge Building, Utterances, Collaborative, Participants, Communicative Functions

1. INTRODUCTION

On-line collaborative knowledge building (OCKB) is assuming an increased role in higher education. In this practice of dialogue supported by chat-room technology, groups of people with shared goals, share activities and experiences in the context of the practice (Nistor, 2012; Lave, 1991; Wenger, 1999). Participation in OCKBs is assumed to lead to the accumulation of experience, stimulation of the social construction of knowledge, and the development of expertise (Nistor, 2012; Bereiter, 2002; Boylan, 2010; Engestrom, 2010). Research in this form of participation often draws on a sociocultural perspective, that is one that emphasizes knowledge building through social and communicative processes. Nevertheless, despite this push towards on-line collaboration approach in both academia and in the workplace as a knowledge building enabler, the studies related to this approach are mainly qualitative which maybe missing or could be complimented by some of the strengths presented by quantitative approaches, including communicative activity investigations and testing hypotheses.

Conflicting views can be identified and resolved collaboratively when participants are working as a community of inquirers (Chan, 2012), but this form of participation often leads to problem solving through creative thinking and inquiry, is open-ended and crucially conducted with others in the group. In this process of joint knowledge building using interactions, participants are active rather passive through exploring, transforming, comparing, coordinating and analyzing different ideas (Hennessy, 2020; Elbers, 1996; Mercer, 2000; Rogoff, 1990). Through these interactions, participants usually make influential responses commonly through *elaborating*, *clarifying* and *building* on previous contributions made by themselves and others in the group (Hennessy, 2020; Elbers, 1996; Mercer, 2000; Rogoff, 1990). The quality of these interactions is therefore paramount (Hennessy et al, 2020).

Participation generally involves natural language dialogue, an essential element of OCKB, and dialogue in an OCKB generally generates a cluster of utterances from multiple participants (Traum, 2018). Identifying the speech acts of utterances (from amongst the cluster) is necessary to correctly classify the participant's intentions. According to (Traum, 2018), it remains a challenge in inferring speech acts (or dialogue) from a surface utterance since an utterance may represent more than one speech act according to the context.

Many researchers of classifying utterances in a knowledge building group, focus on the collective, reciprocal, cumulative and purposeful nature of dialogue (Hennessy, 2020; Alexander, 2008; Mortimer, 2003). This perspective (1) includes construction of meaning through pursuing common goals and chained lines of collaborative inquiry in which answers to questions give rise to new questions (Bakhtin, 1981; Kumpulainen, 2010; Wells, 1999), and (2) recognizes the importance of participants testing their ideas against other as a means of promoting participants' argument literacy ("evaluativist" epistemology) (Hennessy, 2020; Wilkinson, 2017).

The communication functions representing the utterances typed by participants in an OCKB are identified as "dialog acts" (Erkens, 2008). For instance, if participant A would like participant B to send off an email, A could use various utterances to achieve this goal, and these utterances would all share the same dialogue function, i.e. fulfil the same dialogue act. *A dialogue act therefore is a unit in the semantic description of communicative behavior produced by a sender and directed at an addressee, specifying how the behavior is intended to influence the context through understanding of the behavior* (Bunt, 2004). The developed coding system that identifies dialog acts with a long history of about twenty years is the [DAC] Dialogue Act Coding system (Erkens, 2008), therefore the main dialog acts used in this paper as independent variables in the proposed regression model of the OCKB, are *argumentative*, *responsive*, *elicitative*, *informative* and *imperative* in keeping with DAC. A sixth independent variable is pre-OCKB assessment grade (*Pre*). The dependent variable is the post-OCKB assessment grade (*Pst*). We begin with an examination of research literature to define these five (5) dialog act variables. Then we define the regression model of the OCKB.

2. DIALOG ACT VARIABLES OF THE OCKB MODEL

2.1 Argumentative (*Arg*)

Erkens (2008), Fraser (1999), and Heeman (1998) describe argumentative as dialogue acts that represent a temporal, causal, or inferential relation between utterances and use conjuncts such as "but", "because" and therefore as a discourse marker. Indicating a line of argumentation or reasoning, Walter (2012) explains that an argumentative utterance is meant to convince someone else and is aimed at proving (or disproving) some claim. Once both parties are committed to a claim, there is no point in arguing any further. Van Boxtel (2000) studied the effects of two different face-to-face collaborative tasks on upper secondary level students' interactions and found that students who scored high on a pretest formulated more arguments during interaction. In this sense, we build our quantitative regression model of the OCKB by regarding the argumentative dialog act as a contributory factor of positive learning.

2.2 Responsive (*Res*)

As part of a constructive argument between two parties, an explanation is usually followed by either a positive response or a negative response (Erkens, 2008; Walton, 2012). Erkens (2008) and Louwerse (2003) went on to argue that responsive dialog acts (e.g. confirmations, denials and answers) have a backward-looking relation to an earlier utterance while the other four communicative functions are forward looking and give new information. Responsive utterances (such as "right", "no", "yeah", "yep" and "uh-huh") react or refer to preceding utterances. (Stolcke, 2000) stresses that detecting these utterances (as answers) are also semantically significant since they are likely to contain new information. Agreement responses (such as "that's ok","right" and "okay") are verbal responses which minimally show that the speaker has heard (and often understood and accepted) the move to which it responds (Louwerse, 2006). In (Ribeiro, 2019), the automatic multilingual dialog act recognition study showed that agreements were among the most frequent set of dialogs in the datasets collected. We therefore regard responsive dialog acts as an influencing variable in our quantitative regression model of the OCKB.

ISBN: 978-989-8704-22-1 © 2020

2.3 Elicitative (*Eli*)

Elicitative dialog acts (such as questions or proposals) request a response from the dialogue partners and consists of proposal to act or a question for information (Erkens, 2008; Graesser, 1994; Lehnert 1978). Some of the utterances requiring or soliciting a response include “Agree?”, “ ….or….?”, and “Why?”. Stolcke (2000) took additional interest in the backchannel or elicitative short utterance “uh-huh”. These utterances play discourse-structuring roles, e.g., indicating that the speaker should go on talking. These are usually referred to in the conversation analysis literature as “continuers” and have been studied extensively (Jefferson 1984; Schegloff 1982; Yngve 1970). Therefore elicitative dialog acts are included in our quantitative regression model of the OCKB.

2.4 Informative (*Inf*)

Informative dialog acts are statements or opinions transmitting new information. The utterances signal that the speaker intends the hearer(s) to acquire the given information (Eugenio, 2010). Eugenio went further to suggest statements (e.g. “that result makes no sense”, “but household items are not supposed to be taxed”) describe a psychological state of the speaker. Opinions often include such hedges as “I think”, “I believe”, “It seems”, and “I mean”. In the Stolcke (2008) study, it was shown that 26% of all statements elicit backchannels. In Eugenio (2010) it was disclosed that statements, across all three datasets, were the most frequent utterances. In this sense, we build our quantitative regression model of the OCKB by regarding the informative dialog act as a contributory factor of positive learning.

2.5 Imperative (*Imp*)

Imperative dialog acts (command utterances) request an action to be fulfilled by the dialog partner or other group members (Erkens, 2008). For example: “Just give me a call tomorrow by five.”, “What was that again?”, “Please say yes or no.” and “Let me check. Just a second please.”. Yu (2019) suggest commands are among the most crucial utterances as they direct the conversation towards a specific topic, and paired with answer utterances [(positive answer, command) and (negative answer, command)], are the most frequent co-occur tags in the conversations studied. Yu went on further to show that a (negative answer, command) pair usually implies an implicit request to change the topic. In (Mezza, 2018) (Louwerse, 2006) (Anderson, 1991), a map-task challenge in the form of a task-based dialogue corpus was collected. The dialogues involve two participants, one with an empty map and one with a route marked map which must instruct the other speaker to draw the same route. The corpus (collection of texts) was chosen for the study due to its abundance of command utterances (more than 30% of the overall corpus). We therefore include imperative dialog acts in our quantitative regression model of the OCKB.

3. DERIVING A QUANTITATIVE REGRESSION MODEL OF THE OCKB

$Pst = Pre + \text{Dialog Act}_g + \varepsilon$ (1)
$Pst = Pre + \text{Dialog Act}_{g1} + \text{Dialog Act}_{g2} + \varepsilon$ (2)

where $\text{Dialog Act}_{g|g(n)} = \text{Total Weight}_g$ comes from [1-*Arg*|2-*Res*|3-*Eli*|4-*Inf*|5-*Imp*]

Two versions of the quantitative OCKB model are proposed- version 1 and version 2. According to the first version of the model, a participant’s performance on the post OCKB assessment is statistically explained by his/her pre OCKB assessment grade together with the corresponding total weight factor of dialog act_g considering the error ε. A total weight factor represents the total weight of the contribution made by the participant to the corpus for dialog act_g where g={1-*Arg*,2-*Res*,3-*Eli*,4-*Inf*,5-*Imp*}.

According to this second version of the model, a participant’s performance on the post OCKB assessment is statistically explained by his/her pre OCKB assessment grade together with the corresponding total weight factors of dialog act_{g1} and dialog act_{g2} together considering the error ε.

4. AIM OF THE STUDY AND HYPOTHESIS

In correspondence with the quantitative model of the OCKB (versions 1 and 2), our research aims at validating the hypothesized quantitative OCKB model (one or both versions), which implies measuring the respective dialog act variables and verifying the following hypotheses:

$H_1$1. Argumentative dialog acts have an influence on the participant's performance.
$H_1$2. Responsive dialog acts have an influence on the participant's performance.
$H_1$3. Elicitative dialog acts have an influence on the participant's performance.
$H_1$4. Informative dialog acts have an influence on the participant's performance.
$H_1$5. Imperative dialog acts have an influence on the participant's performance.
$H_2$1. Argumentative dialog acts together with informative dialog acts have an influence on the participant's performance.
$H_2$2. Elicitative dialog acts together with imperative dialog acts have an influence on the participant's performance.
$H_2$3. Responsive dialog acts together with imperative dialog acts have an influence on the participant's performance [Yu (2019)].

5. METHODOLOGY

5.1 Design

Data from one study was used. During the study, participants collaborated in small groups in an online chatroom environment on query writing tasks for the computer science related subject of Database Management. As part of the study (1) participants undertook an assessment [Task 1 of 3] before the collaboration exercise [Task 2 of 3] ahead of the second assessment [Task 3 of 3], and (2) the collaborative process between the participants was captured in log files. These log files were manually coded (using the coding scheme in 5.7) for the calculations of the dialog act variable weights (see section 5.8). In terms of quantitative learning improvement through knowledge building as hypothesized (H1 – H3), we investigated the dialog act weights in combination that influenced the second assessment scores given the first assessment scores using a regression model.

5.2 Participants

Participants were second and third year computer science degree students in the Faculty of Engineering and Computing at the University of Technology, Jamaica in the capital city of Kingston, Jamaica. Both cohorts in full were invited to participate from which a total of 59 students accepted. From that group, 44 completed the mandatory three (3) tasks. Ranging from the age of 17 years to 28 years, this final group comprised of 24 males and 20 females. None of the participants had undertaken a formal course in database query writing prior.

5.3 Learning Objectives

The topic of the 3rd year level database management course is called SQL (Structured Query Language), a technique used to query a set of related data tables. At the end of the learning activities over the four (4) consecutive days, participants should be able to write efficient SQL scrips to solve complex queries.

5.4 Task 1

Day 1, 5 hours: Participants were asked to complete a five (5) question assessment individually given a set of related data tables. Each of the questions comprise of a complex query to solve with an SQL script.

5.5 Task 2

Days 2-3: Participants, in groups of 4-5 persons formed by themselves, collaborated exclusively using WhatsApp. WhatsApp is a text messaging, image sharing, video sharing and voice messaging app for mobile devices. With an additional feature of group chatting, this free service facilitated collaborative learning. Participants were given material on SQL to include a different set of related data tables and nine (9) complex query questions. Four of questions were already solved. In keeping with an evaluativist epistemology approach, groups were encouraged to brainstorm inside the group, ask each other probing questions, test each other's SQL scripts and ideas in solving the remaining five questions together. I made myself available to the groups to provide scaffolding support and at strategic points, posted tips from which the groups could generate news ideas.

5.6 Task 3

Day 4, 80 minutes: Participants were asked to complete a five (5) question assessment individually. In fact, this assessment was exactly the same given for the pretest (in Task 1).

5.7 Dialog Act Coding Scheme

The text data collected in Task 2 represented the individual group conversations thus constituted a corpus of participant collaboration. Table 1 shows an example how the DAC was used to code a fragment of four participants collaborating on the SQL script writing task in the same group. The derived functions (or dialog act variables) are counted towards the calculations of the weight factors (described in section 5.8)

Table 1. Example of a coded online collaboration protocol

Participant	**Message**	**Dialog Act Code**	**Description**	**Function/Dialog Act Variable**
302	Inner joins can solve it but I'm not familiar with them	InfEvlNeu	Neutral Evaluation	Inf
309	What ya talking abt	EliQstOpen	Open Question	Eli
315	We solved it I think	InfStmAct	Announcement of actions	Inf
307	Aite I'm done with this practice	InfStmAct	Announcement of actions	Inf
307	I solved them last night	InfStmAct	Announcement of actions	Inf
302	Explain. Pls	ImpAct	Order for action	Imp

5.8 Dialog Act Variable Weight Factor Calculations

Dialog act variable weight factors were calculated to be a representation of respective contribution level percentages of each participant in his/her group. Corpuses were sampled upto the first 60% per group.

n = size of the group corpus sampled (number of text conversations)

Arg_{g} = number of arguments (*Arg*) made by the participant
$Arg_{\mathrm{p}} = Arg_g/n$ [*Arg* weighted factor for the participant]
Res_{g} = number of responses (*Res*) made by the participant
$Res_{\mathrm{p}} = Res_g/n$ [*Res* weighted factor for the participant]
Eli_{g} = number of elicitative (*Eli*) dialog acts made by the participant
$Eli_{\mathrm{p}} = Eli_g/n$ [*Eli* weighted factor for the participant]
Inf_{g} = number of informative (*Inf*) dialog acts made by the participant

$Inf_p = Inf_g/n$ [*Inf* weighted factor for the participant]
Imp_g = number of imperative (*Imp*) dialog acts made by the participant
$Imp_p = Imp_g/n$ [*Imp* weighted factor for the participant]

5.9 Regression

A regression analysis on the data was performed to indicate how effective the learning improvements were. The following hypotheses were tested:

H_11: $Pst = Pre + Arg + \varepsilon$
H_21: $Pst = Pre + Arg + Inf + \varepsilon$
H_14: $Pst = Pre + Inf + \varepsilon$
H_15: $Pst = Pre + Imp + \varepsilon$
H_13: $Pst = Pre + Eli + \varepsilon$
H_23: $Pst = Pre + Res + Imp + \varepsilon$
H_12: $Pst = Pre + Res + \varepsilon$
H_22: $Pst = Pre + Eli + Imp + \varepsilon$

6. RESULTS

Table 2. Regression results

	Variable	**Coefficient**	**t Stat**	**p-Value**	r^2	**Standard Error**	**Based on the p-Value**
H_11					0.23449	12.42861	*Rejected*
	Pre	0.65432	3.40989	0.00155			Sig
	Arg	0.90019	0.75664	0.45393			Not Sig
H_12					0.23849	12.39605	*Rejected*
	Pre	0.66201	3.44573	0.00140			Sig
	Res	0.64477	0.88055	0.38407			Not Sig
H_13					0.21807	12.24309	*Rejected*
	Pre	0.55457	2.86621	0.00673			Sig
	Eli	0.76571	1.32300	0.19374			Not Sig
H_14					0.22712	12.48829	*Rejected*
	Pre	0.63279	3.33412	0.00192			Sig
	Inf	-0.18805	-0.45247	0.65351			Not Sig
H_15					0.31792	11.73181	***Accepted***
	Pre	0.47698	2.51815	0.01613			Sig
	Imp	1.64957	2.30011	0.02703			**Sig**
H_21					0.24659	12.49546	*Rejected*
	Pre	0.67582	3.46707	0.00135			Good
	Arg	1.25196	0.97796	0.33444			Not Sig
	Inf	-0.34317	-0.77104	0.44558			Not Sig
H_22					0.31815	11.88731	*Rejected*
	Pre	0.47856	2.48664	0.01753			Sig
	Eli	-0.08113	-0.11118	0.91207			Not Sig
	Imp	1.71653	1.81898	0.07702			Not Sig
H_23					0.26804	11.84548	***Rejected***
	Pre	0.50354	2.54491	0.01523			Sig
	Res	0.37239	0.52367	0.60362			Not Sig
	Imp	1.58086	2.14815	0.03832			**Sig**

As shown in Table 2, regression model version one revealed that there was no strong statistical evidence to support H_11, H_12, H_13 and H_14. Communicative function *informative* with a p-value of 0.654 ($r^2 = 0.23$) showing the weakest down to *elicitative* at 0.94 ($r^2 = 0.22$). The model (version 1) however, showed a statistical significant evidence to support H_15 with a p-value of 0.027 and r^2 of 0.32. This further suggest that for every percentage gain on the second assessment, an average of 1.65 (coefficient) *imperative* dialog acts was helpful.

For regression model version two also shown in Table 2, it was revealed that the most statistically significant communicative function p-value was 0.04 (*imperative*) to the least p-value of 0.912 (*elicitative*), hence not enough evidence to support $H_2 1$, $H_2 2$ and $H_2 3$. The best r^2 from the version two model was 0.32 ($H_2 2$). Again however, communicative function *imperative* with p-values of 0.03 ($H_2 3$) and 0.77 ($H_2 2$), some evidence was shown to exist statistically that the function had some effect on the second assessment results in combination with another communicative function.

7. DISCUSSION & IMPLICATIONS

The purpose of this study was to validate a proposed dual quantitative model of an OCKB and to investigate (1) the knowledge building effects, if any, of the natural communicative functions between participants in small groups in a computer science related discipline, and (2) by how much these communicative functions individually and in combination, contribute to predicting participant assessment performances.

The results of the regression analysis didn't provide strong enough evidence for the hypothesized dual quantitative OCKB model for seven of the eight tests. The analysis failed to demonstrate statistically that all of communicative functions individually and in combination except those involving *imperative*, had an effect on the final assessment grades thus not supporting $H_1 1$, $H_1 2$, $H_1 3$, $H_1 4$, $H_2 1$, $H_2 2$ and $H_2 3$. The analysis therefore suggests that dialog acts that are *argumentative*, *responsive*, *elicitative* and *informative*, may not be effective towards a positive learning outcome during OCKB sessions. The analysis however provided a statistical significant support for $H_1 5$ suggesting that communicative function *imperative* has some effect on the learning outcome during OCKB sessions. Even in combination with *responsive* utterances as highlighted by (Yu, 2009) [$H_2 3$] and *elicitative* utterances [$H_2 2$] during OCKB sessions, a fairly strong statistical effect of the *imperative* dialog act was demonstrated. The findings therefore suggest that of the five communicative functions, *imperative* contributed most to knowledge building during on-line collaborative learning sessions.

A subsequent examination of some of the group corpuses was conducted to review fragments of the *imperative* utterances used and to learn to whom these utterances were attributed. Figure 3 (below) is an example:

Participant	**Message**	
453	Let me see your code Gregory (*Participant 556*)?	
453	Y'all are moving too quickly	
634	Oh sorry	
556	mek we move through together	[*mek we* means *let us*]

Figure 3. Fragment of mostly imperative utterances from a group corpus

According to Yu (2009), imperative or command utterances (as demonstrated in Figure 3), are used to control the direction of a group discussion which can change at any moment especially if one of the members doesn't find the current conversation useful or appears to be negative. Command utterances are also used to set the mood of the conversation. Participants #453 & #556 in my face to face classes are considered by some of the others in the class as *authoritative* and usually dominate the conversations. Together, both were responsible for 92.50% of the *imperative* utterances made in their OCKB group.

This has implications for the blending in of or reliance on on-line collaborative learning (OLCL) to supplement campus-based or on-line courses. In a South Korean study (Lee, 2017), it was concluded that in face to face classes where the teacher (or lead) has gotten students' approval, the teacher (or lead) has an implicit form of power in the form of an authority which allows for the class lessons to be delivered more effectively. Lee stressed that approval is not given until the teacher (or lead) has demonstrated certain attributes or qualities: demonstration of qualifications, a strong background in content knowledge, demonstrated enthusiasm, and the ability to understand students. But this form of authoritative discourse according to Bakhtin (1981), is fixed in a learning setting, that gives the learners little or no space to personalize meaning or information received by inserting other meanings situated in their context. Therefore, in a group or classroom, learners or participants are forced to either uphold or disregard the discourse. Every learning group has or develops an implicit hierarchical social structure (Lee, 2017) that tends to be driven by or allows some form of authoritative discourse as a result, but for authoritative discourse to work in OLCL sessions and for the lead person(s) to be accepted and be approved by the other participants in the learning group, Lee (2017) suggests relationships between the group members in cultural meaningful ways be developed early, and that the lead takes the time to learn more about the other group members, their contexts of living and to respect their respective cultures and beliefs.

This research aimed at proposing a quantitative model of the OCKB model that describes or better statistically explain the relationship between dialog act variables and a positive learning outcome among the participants in the collaboration. Results from the research suggest this quantitative approach along with some explanatory qualitative analysis, provides a powerful framework for understanding OLCL. The importance of the imperative function was demonstrated necessary for knowledge building, however the function might be a product of influence and power dimensions which underscores interactions. Understanding the strong possibility of influence that can be generated during interactions, the OCKB model can provide valuable insights to teachers and researchers in understanding group dynamics during an OLCL.

This study has some limitations. First, while manual coding of the utterances may guard against coding in isolation, considering preceding or following utterances, the final codings are subjected to inconsistencies especially because analyzing interactions often requires an interpretation of the context within which the utterance was used (Erkens et al, 2008). Future research in this methodological issue in addressing reliability of the coding with some level of automation is needed. Second, the rejection of so many of the hypotheses, suggests a level of weakness in the proposed OCKB models. A key contributory factor is the fact that the dialog weight factor calculations used weren't founded in literature. Future research should seek to strengthen and validate these calculations.

The results of the study provide some evidence that knowledge building during OLCL sessions can be quantified statistically from the corpuses perhaps only from the *imperative* communicative functions as the proposed model proved. Instructors, in allowing students or participants to form groups to collaborate for learning, should be mindful of the hierarchical social structure that forms implicitly and the possible learning impact on the respective groups from the structures.

ACKNOWLEDGMENT

This research was undertaken as part of the PhD in E-research and Technology Enhanced Learning in the Department of Educational Research at Lancaster University. I am pleased to acknowledge the contribution of tutors and peers in supporting the development of this study and its report as an assignment paper.

REFERENCES

Alexander, R. J. (2008). Towards dialogic teaching. Rethinking classroom talk (4th Ed). Cambridge: Dialogos, UK.

Bakhtin, M. M. (1981). The dialogic imagination: Four essays (M. Holquist, Ed.; C. Emerson & M. Holquist, trans.). Austin, TX: University of Texas Press.

Bereiter, C. (2002). Education and mind in the knowledge age. Mahwah, NJ: Lawrence Erlbaum.

Bex, F., Walton, D. (2012) Combining Explanation and Argumentation in Dialogue. ECAI 2012 Workshop on Computational Models of Natural Argument (CMNA), Montpellier, France.

BoyIan, M. (2010). Ecologies of participation in school classrooms. Teaching and Teacher Education.

Bunt, H. (2004). A Framework for Dialogue Act Specification.

Chan, C., Lam, I., Leung, R. (2012). Can Collaborative Knowledge Building Promote Both Scientific Processes and Science Achievement? International Journal of Educational Psychology, 1(3), 1 99-227.

Elbers, E. (1996). Cooperation and social context in adult-child interaction. Learning and Instruction.

Engestrom, Y., Sannino, A. (2010). Studies of expansize learning: Foundations, findings and future challenges. Educational Research Review.

Erkens, G., Janssen, J. (2008). Automatic coding of dialogue acts in collaboration protocols. International Journal of Computer-Supported Collaborative Learning.

Fraser, B. (1999). What are discourse markers? Journal of pragmatics, 31, 931-952.

Graesser, A. C., Person, N. K., (1994). Question asking during tutoring. American Educational Research Journal, 31, 104-137.

Heeman, P. A., Byron, D., Allen, J. F. (1998). Identifying discourse markers in spoken dialog. AAAI Spring Symposium on Applying Machine Learning & Discourse Processing.

Hennessy, S., Howe, C., Mercer, N., Vrikki, M. (2020). Coding classroom dialogue: Methodological considerations for researchers. Learning Culture and Social Interaction.

Jefferson, G. (1984). "Notes on some orderlinesses of overlap onset". In V. D'Urso & P. Leonardi (Eds.), Discourse Analysis and Natural Rhetorics (pp. 11–38). Padua: Cleup Editore.

Jeong-A, L., Chan-Jong, K. (2017). Teaching and Learning Science in Authoritative Classrooms: Teachers' Power and Students' Approval in Korean Elementary Classrooms. Res Sci Education.

Kumpulainen, K., Lipponen, L. (2010). Productive interaction as agentic participation in dialogic enquiry. Educational Dialogues: Understanding & Promoting Productive interaction. Abingdon, Oxon: Routeledge.

Lave, J., Wenger, E (1991). Situated Learning. Legitimate peripheral participation. Cambridge, UK: University Press

Lehnert, W. G. (1978). The process of question answering: A computer simulation of congnition. Hillsdale, NJ: Lawrence Erlbaum Associates.

Louwerse, M., Crossley, S. (2003). Dialog Act Classification Using N-Gram Algorithms.

Mercer, N. (2000). Development through dialogue. In N. Mercer (Ed.), words and minds: How we use language to think together.

Mortimer, E. F., Scott, P. H. (2003). Meaning making in secondary science classrooms. Maidenhead: Open University Press.

Nistor, N., Fischer, F. (2012). Communities of practice in academia: Testing a quantitative model. Learning, Culture and Social Interaction.

Ribeiro, E., Ribeiro, R., Martins de Matos, D. (2019). A Multilingual and Multidomain Study on Dialog Act Recognition Using Character-Level Tokenization. MDPI.

Rogoff, B. (1990). Apprenticeship in thinking: Cognitive development in social context, Oxford: Oxford University Press

Stolcke, A., Coccaro, N., Bates, R. (2000). Dialogue Act Modeling for Automatic Tagging and Recognition of Conversational Speech. Association for Computational Linguistics.

Traum, D., Henry, C., Lukin, S., Artstein, R. (2018). Dialogue Structure Annonation for Multi-Floor Interaction.Proceedings of the Eleventh International Conference on Language Resources and Evaluation.

Van Boxtel, C., Van der Linden, J., Kanselaar, G. (2000). Collaborative learning tasks and the elaboration of conceptual knowledge. Learning and Instruction, 10, 311-330.

Wells, G. (1999). Dialogic Inquiry: Towards a socioculturalpractice and theory of education. Cambridge: Cambridge University Press.

Wenger, E. (1999). Communities of practice. Learning, meaning and identity. Cambridge, UK: University Press.

Wilkinson, I. A. G, Reznitskaya, A., Bourdage, K. (2017). Towards a more dialogic pedagogy: Changing teachers' beliefs and practices professional development in language arts classrooms. Language and Education.

Yu, D., Yu, Z. (2019). MIDAS, A Dialog Act Annotation Scheme for Open-Domain Human-Machine Spoken Conversations. arXiv.org