Working Paper No. 2009-1
February 2009

Instituting Change in Classroom Discourse Structure: Human and Computer-Based Motif Analysis

Mitchell J. Nathan
Department of Educational Psychology/
Wisconsin Center for Education Research
University of Wisconsin–Madison
mnathan@wisc.edu

Suyeon Kim
Wisconsin Center for Education Research
University of Wisconsin–Madison
kim14@wisc.edu

Timothy S. Grant
Department of Educational Psychology
University of Wisconsin–Madison
tsgrant@wisc.edu

WCER
Wisconsin Center for Education Research
School of Education • University of Wisconsin–Madison • http://www.wcer.wisc.edu/

WCER working papers are available on the Internet at http://www.wcer.wisc.edu/publications/workingPapers/index.php. Recommended citation:

Nathan, M. J., Kim, S., & Grant, T. S. (2009). *Instituting change in classroom discourse structure: Human and computer-based motif analysis* (WCER Working Paper No. 2009-1). Madison: University of Wisconsin–Madison, Wisconsin Center for Education Research. Retrieved [e.g., February 15, 2008,] from http://www.wcer.wisc.edu/publications/workingPapers/papers.php

The research reported in this paper was supported in part by the U.S. Department of Education Institute of Education Sciences under Grant No. R305H060097 and by the Wisconsin Center for Education Research, School of Education, University of Wisconsin–Madison. Any opinions, findings, or conclusions expressed in this paper are those of the authors and do not necessarily reflect the views of the funding agencies, WCER, or cooperating institutions.

Instituting Change in Classroom Discourse Structure: Human and Computer-Based Motif Analysis

Mitchell J. Nathan, Suyeon Kim, and Timothy S. Grant

When people talk, they tend to adopt regular patterns of conversational organization. What forces influence the organizational structures of classroom discourse? In this paper, we report on a study comparing the structure of classroom discussions in a middle school mathematics classroom before and after teacher participation in professional development activities aimed at enhancing students' classroom participation and the co-construction of mathematical ideas. We show that changes in the classroom were accompanied by identifiable changes in discourse structure. In particular, the de-centering of the teacher's mathematical authority led to a reduction in traditional teacher-led *initiation-response-evaluation* (IRE) patterns, and to increases in student-led *initiation-demonstration-evaluation* (IDE) patterns.

Theoretical Framework

Monologic discourse focuses on response to an authority in an expected manner, often as a way of showing adherence to the canon (Bakhtin, 1986; Hakkarainen & Paavola, in press). *Dialogic discourse*, in contrast, derives from a *participatory* view of learning (Sfard, 1998, 2008) that frames knowledge as distributed and culturally bound, emphasizing the socially mediated nature of distributed, personalized knowledge generation. Dialogic discourse is a powerful way to promote student engagement and higher order reasoning (Nathan & Kim, in press), long-term retention, and transfer of concepts to new contexts. To transform monologic instruction into dialogic interactions within the classroom, the locus of authority of knowledge must be de-centered, and students granted permission to initiate discussion, have protracted turns-at-talk in which they explore their ideas, and legitimately provide evaluations of the accuracy and appropriateness of a peer's contributions.

Everyday talk and institutional talk. The everyday talk of casual conversation and domestic affairs is subject to simple organizational constraints (Sacks, Schegloff, & Jefferson, 1974). Institutional talk, as found in the workplace and the classroom, differs from everyday talk (e.g., Drew & Heritage, 1992) because those in positions of authority have ethical and professional obligations constraining their prompts and responses. Patterns such as IRE sequences are pervasive during pedagogical exchanges (Greenleaf & Freedman, 1993; Mehan, 1979; Sinclair & Coulthard, 1975) and often dominate educational discourse (e.g., Lemke, 1990; Wells, 1993). In a typical IRE pattern, the teacher initiates the exchange (an *I*-event), generally by asking a closed, "known-answer" question (*display question*). This elicits a direct response (*R*-event) from a student, which is then evaluated (*E*-event) by the teacher, often in a way that terminates the interaction.

Evaluative and non-evaluative exchanges. Nunan (1987) and others (e.g., Lemke, 1990; Nystrand, 1997; Thornbury, 1996; Wood, 1992) have been critical of IRE exchanges because they offer students limited opportunities to verbalize their own ideas. Cullen (2002) distinguished between the evaluative and the discoursal roles of the third element of the teacher-student IRE interaction. In the evaluative phase, teacher feedback is given to directly accept or reject the student response, situating judgment within the context of a power imbalance and often failing to elicit further participation (cf. Nystrand, 1997). Thus, we observe a monologic exchange (Bakhtin, 1986; Lotman, 1988) that may be useful in delivering information and assessing knowledge but does little to elicit multiple perspectives.

When a non-evaluative follow-up question replaces the evaluation phase (i.e., *initiation-response-follow-up*, or IRF), it provides more impetus to perpetuate the discourse (Wells, 1993; Wells & Arauz, 2006). In IRF exchanges, the follow-up movement, or *F-movement* (Sinclair & Coulthard, 1975; Cullen, 2002), invites multiple perspectives and typically leads to further contributions from students, thus sustaining their involvement and increasing both their engagement (Mercer, 1995) and their participation in the kinds of open-ended exchanges that can promote higher order reasoning (Fernyhough, 1996).

Student agency in classroom discourse. In the context of contemporary education reform aimed at promoting learner-centered classrooms and students' construction of knowledge, it is more common to see classrooms with students initiating discursive sequences and acting as the primary agents for enacting evaluation and follow-up events (Engle & Conant, 2002; Lampert, 1990; Stipek et al., 1998). These interactions may elicit protracted demonstrations of students' ideas that are far more elaborate and personalized than typical R-events that follow known-answer questions. Sequences built around these demonstrations, or D-events, can dominate highly participatory classrooms and show frequent chaining. For example, Nathan, Eilam, and Kim (2007) documented IDE sequences in 77.8% of the exchanges of a whole-class collaborative problem-solving session, with chaining from one IDE sequence to the next evident in 81% of the exchanges. They suggested that the precipitating influences of the IDE structure might be found in the established norms of interaction and the creation of a respectful and secure classroom environment, which enable students to dominate the discussion and serve as principal agents for evaluating and setting the direction of the discourse.

Hypotheses

In our study of the structure of discourse in a middle school mathematics classroom, we hypothesized that:

1. Traditional IRE/F patterns would predominate in the first year; and

2. IRE/F patterns would decrease and IDE patterns increase in the second year following teacher professional development aimed at enhancing students' participation and the co-construction of mathematical ideas.

Method

Participants and Setting

We observed and videotaped lessons from a sixth-grade public school mathematics class at various points over a 2-year period. In each of the 2 years, the students exhibited a wide range of mathematical performance (standardized tests from 5^{th} to the 99^{th} percentile). The teacher, whose training was in elementary education, was nominated by her principal to participate. She expressed a strong professional commitment to develop her classroom discourse techniques so that she could implement the kinds of instruction called for in reform documents such as the

National Council of Teachers of Mathematics' (1989, 1991, 2000) principles and standards for curriculum, evaluation, and teaching.

For this study, we selected two lessons each from Years 1 and 2 based on the condition criterion that each contain an extended, whole-classroom discussion about mathematics that was consistent with the teacher's planned curriculum and therefore constituted ordinary instructional time.

Professional Development Intervention

Over a 3-week period during the summer between Years 1 and 2, the teacher participated in one-on-one conversations with our professional development team about reform instruction, her professional and personal goals as a teacher, and video-prompted reviews of her Year 1 classroom teaching and that of others (e.g., Ball, 1993). The summer program included readings from math teachers and education researchers about (a) learning, curriculum, and participatory forms of instruction; (b) active listening and other methods of achieving classroom norms conducive to student participation and student-directed learning; (c) ways of presenting ideas to the classroom; (d) approaches to modeling constructive feedback; and (e) techniques for eliciting and using students' multiple solution methods to facilitate mathematical participation and learning.

Coding and Pattern Identification

The data set consisted of four lessons, two from each year. The classroom discourse was captured on video and digitized, then imported into Transana, a computer application for discourse analysis (Fassnacht & Woods, 2005; www.Transana.org). Each videotape was transcribed, segmented into analytic units, and then further subdivided into coded events—such as I-, R-, D- and E-events and F-movements—consistent with prior work reported by others (Lemke, 1990; Mehan, 1979; Nathan et al., 2007; Sinclair & Coulthard, 1975; see Appendix A for codes and examples). This data was then subjected to both human and computer analysis, as described below. Units in the transcript were bracketed by video time codes that made it possible for movement through one medium (e.g., video) to be coordinated with movement through the other (e.g., transcript). Transana visually highlights the corresponding region of the transcript during video playback, which allowed us to easily link the speech with the videotaped actions, and vice versa.

Pattern identification using human coders. Once an entire lesson was transcribed, segmented, and coded by members of the research team, the stream of codes was examined for recurrent patterns that would indicate organizational structure within the discourse. Previously documented patterns such as IRE, IRF, and IDE provided theoretical guidance for the pattern-finding process, though coders were open to finding other patterns as well. Pattern finding consisted of iterative examination of the string of codes (Appendix B), the transcript, and the video using the integration supported by Transana.

Pattern identification using computer-based motif analysis. Motif analysis uses random walks based on probabilities to find likely patterns and potential starting sites within coded sequences (Cadez, Heckerman, Meek, Smyth, & White, 2003; Grant, 2007; Keles, van der Laan,

& Eisen, 2002). For example, a portion of the code stream from Lesson 1 from Year 1 of our corpus is shown in Figure C1 (Appendix C). The algorithm begins by segmenting the data stream of temporally ordered codes from the transcript (10 codes in each segment; see b. in Figure C1), then establishing a random starting point (c. in Figure C1). Each segment is assigned a motif at random (e.g., Dg-EE-Ti; c. in Figure C1) and a site (fourth position) at which the motif starts. Starting motifs (d. in Figure C1) are defined based on the segments that are assigned a common motif. Note that in the example the algorithm is considering a DEI pattern, which, when part of a chain, may be an IDE motif (where the final I in DEI is the starting position in the subsequent IDE sequence).

Figure C2 illustrates the random walk of the motif algorithm as it identifies patterns. A selected segment is identified by the particular motif and the site at which the motif begins. The motifs are all defined based on these assignments. First, a segment is selected (Step 1). Based on the current definitions of motifs (Step 2), new motifs and starting positions are selected at random, with probabilities defined by how likely each motif is to start at each position (Step 3). Then, the motifs are redefined based on the new assignments of motif and starting position (Step 4). This process is repeated thousands of times. Over the course of the random walk, the definitions of motifs and the assignment of motifs will converge to within a region of probable (not optimal) solutions.

Results

As noted earlier, we hypothesized that the coded sequences would reveal the presence of IRE/F and IDE patterns and that Year 2 would show a greater presence of IDE patterns as a result of changes in the role of the teacher and students in the classroom.

Pattern Identification by Human Coders

As Table 1 shows, in Year 1 IRE and IRF patterns dominated (IRE: n = 25; IRF: n = 15), and IDE sequences were extremely rare (n = 1). The remaining five event sequences did not fit any preexisting category that had been identified in the prior research literature on classroom discourse.

Table 1
Frequency of IDE, IRE, and IRF Patterns Identified by Human Coders for the Lessons in Years 1 and 2

	IDE	IRE	IRF	No Code
Y1-1	1	20	11	4
Y1-2	0	5	4	1
Y2-1	7	3	7	2
Y2-2	11	6	4	2
Total	19	34	26	9

In Year 2, the pattern identification showed a marked change in the discourse structure. Consistent with our predictions about the influences of a shift in the teacher's approach to discourse in the classroom, IDE pattern use increased (n = 18). IRE and IRF patterns, while still prevalent, occupied a smaller portion of the discourse (IRE: n = 9; IRF: n = 11).

Following scholars such as Wells (1993), we combined IRE and IRF patterns since both featured teacher-initiated discussion, short student responses, and teacher-led evaluations or follow-ups. A chi-square analysis on frequency of IRE/F vs. IDE patterns in Years 1 and 2 was significant, chi-square (1) = 33.5, $p < .01$. Consistent with our hypotheses, there was a greater presence of IRE/F patterns in the first year than in the second year, which saw an increased presence of the IDE pattern.

Pattern Identification Through Computer-Based Motif Analysis

Exploratory analyses. The human identification of the event sequences suggested that there were identifiable patterns within the discourse. However, as is evident from a cursory inspection of Appendix B, there were many complexities to the data—including event sequence variants, insertions, and partial sequences—that made the human identification of patterns demanding, fraught with subjectivity, and potentially of low reliability. We used the motif analysis in the exploratory, or data-driven, mode for motifs of length 3 (to find IRE, IRF, IDE) and length 4 (to provide context for these patterns) (Keles et al., 2002).

For Year 1, Lesson 1, the dominant motif that emerged with a window of length 3 was an IRE pattern, evident in 20 of the 23 segments (Table 2). When a window of length 4 was used, the dominant motif was EIRE (21 out of 26 segments; Table 3). This result reflects one of the ways the larger window shows the context of these patterns and the way the algorithm deals with the IRE pattern chaining throughout the lesson, with E-events from a prior IRE sequence serving as the trigger for a subsequent IRE sequence. Year 1, Lesson 2, showed the same dominant motifs for length 3 (Table 2) and length 4 (Table 3), but the algorithm also predicted they would occur less frequently.

Table 2
Results from the Exploratory Motif Analysis: Motif Window of Length 3

		Frequency of				
Session	**Dominant motif (length 3)**	**Dominant motif**	**Alternate motif**	**Alternate motif**	**No motif**	**Total segments**
Y1-1	IRE	20	1	0	2	23
Y1-2	IRE	8	1	1	1	11
Y2-1	IR(IE) or DEI	9	4	1	3	17
Y2-2	IDE	8	4	2	2	16

Table 3
Results from the Exploratory Motif Analysis: Motif Window of Length 4

		Frequency of				
Session	**Dominant motif (length 4)**	**Dominant motif**	**Alternate motif**	**Alternate motif**	**No motif**	**Total segments**
Y1-1	EIRE	21	2	1	2	26
Y1-2	EIRE	7	2	0	3	12
Y2-1	(RE)I(RD) E	13	2	2	3	20
Y2-2	EI(RD)E	14	4	0	1	19

The Year 2 lessons, following the professional development, showed some departure from the IRE pattern, with the emergence of more open-ended initiation prompts, followed by student demonstrations of their own mathematical ideas. In the length 3 motif (Table 2), the dominant motif showed a mix of the IDE and IRE sequences, shown as I(R or D)E patterns, where R- and D-events shared dominance. This was found in 9 of the 17 segments from the first lesson of Year 2 (Table 2). However, as Table 2 shows, there were also strongly competing alternate motifs, specifically DEI and IR(I or E). The algorithm appears to have found a point of multiple equilibria, where no single motif received enough of an advantage to become dominant and segments just moved equally between them. Note that some of the dominant and alternate motifs can be reordered to make an IRE or IDE pattern, which reveals the chaining effect, where an E or D from a prior triad occupies the first position of a subsequent IRE or IDE sequence.

The length 4 motif for Year 2 Lesson 1 (Table 3) showed that multiple codes were substitutable in the first and third sites. Even though this pattern was flexible, it was also strong, predicted to appear in 13 of the 20 segments. The pattern also showed leading evaluation and short-answer response events that may precede (and trigger) IRE and IDE patterns.

Year 2 Lesson 2 clearly showed that the discourse was at times organized around student-led demonstrations in response to initiations, along with a reduction in display (closed) questions posed by the teacher. The dominant motif of length 3 (Table 2) was characterized as an IDE pattern and found in 8 of 16 segments. The longer length 4 motif (Table 3) showed again the leading context of an evaluation followed by an initiation event, and then the substitutability of codes—this time in the third site, where R- and D-events appeared to play a similar role in the discourse structure. Thus, we see in this final lesson the co-dominance of IRE and IDE motifs.

Confirmatory analyses. Four patterns generated by the human analysis and the exploratory motif analysis (length 3) were selected for specific investigation by means of a confirmatory motif analysis: (a) the traditional IRE pattern, (b) an IRE-1 hybrid pattern that includes both teacher-directed initiation and evaluation events as in IRE motifs (see the codes of Table 5) and extended student responses as in IDE motifs, (c) an initiation-response-initiation (IRI) pattern identified by the computer-based motif analysis but not by the human coders, and (d) the IDE pattern. Table 4 shows the frequencies with which these motifs (as defined in Table 5) occurred in the data. As predicted, the IRE pattern was the major organizational structure in the Year 1 lessons, even taking into account variations in lesson length. IRE continued to play an important role in Year 2. Also as predicted, the IDE pattern, which was negligible in Year 1, was clearly evident in Year 2, as shown by both the human and computer coding. The IRI pattern was also an important structure in the classroom discourse across all the lessons and represents an important discovery from the motif analysis.

Table 4
Results of the Confirmatory Motif Analysis: Frequency of Patterns Found (with Motif Window of Length 3)

	Y1-1	Y1-2	Y2-1	Y2-2
Motif IRE	32	16	10	11
Motif IRE-1[a]	6	1	4	8
Motif IRI	17	6	15	7
Motif IDE	1	0	7	11
Number of codes	186	84	138	126

[a]IRE-1 emerged from the computer-based exploratory analysis as a hybrid of IRE and IDE.

Table 5
Results of the Confirmatory Motif Analysis: Codes from Appendix A Used in the Definition of Each Motif

	1st site	2nd site	3rd site
Motif IRE	Ti TI TF	RR	Ee EE NE Te TE
Motif IRE-1[a]	Ti TI TF SI	DD Dg DG DW	Ee EE NE Te TE
Motif IDE	TI SI BI	DD Dg DG DW	Ee EE Te TE Se SE Be BE NE

[a]IRE-1 emerged from the computer-based exploratory analysis as a hybrid of IRE and IDE.

The probabilistic structure of motifs allows us to quantify how unlikely it is that these patterns would occur in their reported frequencies by chance alone. The less likely, the more support we have for the claim that these do actually reflect patterns in the data. For our cutoff, we used occurrences of less than 1 in 10,000 ($p < .0001$). IRE patterns in all four lessons were significant, as were IDE patterns in Year 2 (Table 6). The table also shows that the frequency of the newly discovered hybrid IRE-1 pattern was significant in lesson 1 of Year 1 and lesson 2 of Year 2. The dramatic rise in Year 2 of the IDE pattern is even more striking, since this is a sequence of largely student-directed discourse. This provides further support for the hypothesized changes in discourse structure with the changes in the classroom environment.

Table 6
Results of the Confirmatory Motif Analysis: Probabilities of Observed Occurrences Based on Frequencies of Each Motif (with Motif Window of Length 3)

	Y1-1	**Y1-2**	**Y2-1**	**Y2-2**
Motif IRE	< 0.0001	< 0.0001	< 0.0001	< 0.0001
Motif IRE-1	< 0.0001	0.21	0.0007	< 0.0001
Motif IDE	0.25	1.0000	< 0.0001	< 0.0001
Motif IRI	0.0003	0.0725	0.0001	0.15
Number of codes	186	84	138	126

Discussion

Often, analysts of classroom talk address the important changes in the content of discourse that are evident with the implementation of reform-based teaching practices (e.g., Yackel & Cobb, 1996). We demonstrate that changes in the climate of the classroom that invite greater student participation in mathematical interactions can also lead to identifiable changes in the discourse structure. In Year 2 of the study reported here, the class enacted far more student-directed IDE sequences than in Year 1, though IRE remained an important construct throughout the data. Analyses conducted by human coders were corroborated by computer-based motif analyses that identified flexible patterns using probabilistic, data-mining methods, while also predicting some novel discourse structures, such as the hybrid IRE-1 form. We want to underscore the synergy between the two analytic methods. The motif analysis is really a *computer-assisted* process, because it was necessary to first code the data manually and then move between the numerical output of the program and the human interpretation of the transcript to which the motif sites referred. It is only with these hermeneutic influences that the motifs gain meaning and lend support to the analyses of the classroom interactions.

References

Bakhtin, M. M. (1986). *Speech genres and other late essays* (V. W. McGee, Trans.). Austin, TX: University of Texas Press.

Ball, D. L. (1993). With an eye on the mathematical horizon: Dilemmas of teaching elementary school mathematics. *Elementary School Journal, 93*, 373–397.

Cadez, I., Heckerman, D., Meek, C., Smyth, P., & White, S. (2003). Model-based clustering and visualization of navigation patterns on a web site. *Data Mining and Knowledge Discovery, 7,* 399–424.

Cullen, R. (2002). Supportive teacher talk: The importance of the F-move. *ELT Journal*, *56*(2), 117–127.

Drew, P., & Heritage, J. (1992). *Talk at work: Interaction in institutional settings*. Cambridge, England: Cambridge University Press.

Engle, R. A., & Conant, F. R. (2002). Guiding principles for fostering productive disciplinary engagement: Explaining an emergent argument in a community of learners classroom. *Cognition and Instruction*, *20*(4), 399–483.

Fassnacht, C., & Woods, D. (2005). Transana (Version v2.0x) [Computer software]. Madison: University of Wisconsin–Madison, Wisconsin Center for Education Research. Available from http://www.transana.org

Fernyhough, C. (1996). The dialogic mind: A dialogic approach to the higher mental functions. *New Ideas in Psychology*, *14*, 47–62.

Grant, T. S. (2007). *An implementation of motif data mining on protocol reports for contextual patterns.* Unpublished manuscript, University of Wisconsin–Madison.

Greenleaf, C., & Freedman, S. W. (1993). Linking classroom discourse and classroom content: Following the trail of intellectual work in a writing lesson. *Discourse Processes*, *16* (4), 465–505.

Hakkarainen, K., & Paavola, S. (in press). Toward a trialogical approach to learning. In Schwarz, B., Hershkowitz, R., & Dreyfus, T. (Eds.), *Guided construction of knowledge in classrooms* (Advances in Learning and Instruction). Boston: Elsevier.

Keles, S., van der Laan, M. J., & Eisen, M. B. (2002). Identification of regulatory elements using a feature selection method. *Bioinformatics*, *18*, 1167–1175.

Lampert, M. (1990). Connecting inventions with conventions. In L. P. Steffe & T. Wood (Eds.), *Transforming children's mathematics education* (pp. 253–265). Hillsdale, NJ: Lawrence Erlbaum.

Lemke, J. L. (1990). *Talking science: Language, learning and values*. Norwood, NJ: Ablex.

Lotman, Y. M. (1988). Text within a text. *Soviet Psychology, 26*(3), 32–51.

Mehan, H. (1979). *Learning lessons: Social organization in the classroom*. Cambridge, MA: Harvard University Press.

Mercer, N. (1995). *The guided construction of knowledge: Talk amongst teachers and learners*. Clevedon, Avon, England: Multilingual Matters.

Nathan, M. J., Eilam, B., & Kim, S. (2007). To disagree, we must also agree: How intersubjectivity structures and perpetuates discourse in a mathematics classroom. *Journal of the Learning Sciences*, *16*(4), 525–565.

Nathan, M. J., & Kim, S. (in press). Regulation of teacher elicitations in the mathematics classroom. *Cognition and Instruction*.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.

National Council of Teachers of Mathematics. (1991). *Professional standards for teaching school mathematics*. Reston, VA: Author.

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics.* Reston, VA: Author.

Nunan, D. (1987). Communicative language teaching: Making it work. *ELT Journal, 42*(1), 136–145.

Nystrand, M. (with Gamoran, A., Kachur, R., & Prendergast, C.). (1997). *Opening dialogue: Understanding the dynamics of language and learning in the English classroom*. New York: Teachers College Press.

Sacks, H., Schegloff, E., & Jefferson, G. (1974). A simplest systematics for the organization of turns-taking in conversation. *Language, 50*, 696–735.

Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher, 27*(2), 4–13.

Sfard, A. (2008). *Thinking as communicating: Human development, the growth of discourses, and mathematizing* (Learning in Doing: Social, Cognitive and Computational Perspectives). New York: Cambridge University Press.

Sinclair, J. M., & Coulthard, R. M. (1975). *Towards an analysis of discourse: The English used by teachers and pupils*. London: Oxford University Press.

Stipek, D., Salmon, J. M., Givvin, K. B., Kazemi, E., Saxe, G., & MacGyvers, V. L. (1998). The value (and convergence) of practices suggested by motivation research and promoted by mathematics education reformers. *Journal for Research in Mathematics Education, 29,* 465–488.

Thornbury, S. (1996). Teachers research teacher talk. *ELT Journal, 50*(4), 279–289.

Wells, G. (1993). Reevaluating the IRF sequence: A proposal for the articulation of theories of activity and discourse for the analysis of teaching and learning in the classroom. *Linguistics and Education, 5,* 1–38.

Wells, G., & Arauz, R. M. (2006). Dialogue in the classroom, *Journal of the Learning Sciences, 15,* 379–428.

Wood, D. (1992). Teaching talk. In K. Norman (Ed.), *Thinking voices: The work of the National Oracy Project* (pp. 203–214). London: Hodder & Stoughton.

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education, 27*(4), 458–477.

Appendix A
Event Codes Used, with Examples from Transcripts

Code	Description	Criteria and example
Ti	Teacher's display question	→ T: So what does one represent? One hour? S: Yeah
Si	Student's display question	→ S: Then, what else is blue? S2: Five.
Bi	Teacher's and student's display question	→ S: Do you want us to draw one of the other dots too? → T: Sure, can you find another? S: Um, two fourths.
TI	Teacher's open question	→ T: Would you be willing to show us why you got five? And I'll be interested to see the reasoning.
SI	Student's open question	→ S7: Who would like to speak? → T: Now can I speak?
BI	Teacher's and student's open question	→ S: What's an improper fraction? → T: What does that mean, improper?
RR	Student's short, direct verbal response	T: What color? → S: Yellow.
DD	Student's demonstration with drawing	S: Jones' one hour and combined them together ((coloring one column in a table on the board)), like that one and that → right there ((drawing a new vertical line in one column in the table)).
DG	Student's demonstration with gestures	T: So, I just want her to talk about her technique. → S: So he did that much in an hour, and she did that much in an hour ((pointing to one column with an index finger and pointing to another column at the bottom table on the board)).
Dg	Student's gesture-only demonstration	T: Come up and point for me. → S: ((Pointing one point on the graph on the overhead projector))
DW	Student's demonstration with writing	S: And then, I just write times seven over seven ((writing a formula "7/7" on the board)).
TF	Teacher's F-movement	S: I'm going to add these two. → T: Why are you going to add them?
SF	Student's F-movement	S6: Blue. → S7: How did you get that?

Code	Description	Criteria and example
BF	Teacher’s and student’s F-movement	S2 : I’m thinking that they split the wall in half. → S3 : But why would it, (inaudible) any higher if (inaudible) hours? → T: What do you think about her question, Jane?
Tf	Teacher’s subsequent F-movement	S: Well, like, it’s an odd number so you can’t really have . . . → T (TF): Which is an odd number? S: Seven. → T (Tf): Oh, seven’s an odd number?
TE	Teacher’s valenced evaluation	S: One half. → T: Good.
SE	Student’s valenced evaluation	T: What’s the point of this? S5: To see the number. . . . → S6: No, no, no.
NE	Teacher’s neutral evaluation	S: I added them. → T: Okay.
Te	Teacher’s elaboration	S: If you, if you have a full wall, it’s (inaudible) ten hours. → T: He’s, he’s saying, he is saying that if you add three and seven to get ten, that’s really two walls. That’s Miss, Miss Jones doing a whole wall and Mr. King doing a whole wall.
Se	Student’s elaboration	S1: It’s either two hours or a half, or one-half hour. → S2: Or four hours.
EE	Teacher’s evaluation and valenced elaboration	S: Twenty-one over, twenty-one over twenty-one. → T: Right. Twenty-one over twenty-one would be exactly one. So it’s really close to one.
ee	Student’s valenced evaluation and elaboration	S1: It was yellow. → S2: One, two, three, four, five, yellow. → S3: No. No. It’s because it’s alright, so if the tenth one was yellow. . . .
BE	Teacher’s and student’s valenced evaluation and/or elaboration	T: If you used ten as your . . . as your numerator? John says twenty, yes? S: Yeah. → T: So it’s really close to a half. Isn’t it like really close? → S: If you double it, then twenty twenty ones.

Appendix B
Example Section from Year 1, Lesson 1, of the Stream of Codes Obtained from the Four Lesson Transcripts Used for Both Human and Computer-Based Pattern Finding

Ti *	Ti †	RR	Ti †	Ti †
RR *	RR †	Te	RR †	RR †
TF *	Te †	Bi	TE †	Te †
		RR		
Dg	TI ‡	Te	Bi	Ti †
EE	DG ‡	Ti	RR	RR †
Ti	Te ‡	RR	TF	TE †
DD		TI	RR	
EE	Ti †	RR	Te	Ti †
	RR †	NE		RR †
Ti †	NE †	TI	Ti *	TE †
RR †		RR	RR *	
EE †	Ti	NE	TF *	Bi
	Dg			RR
Ti †	TE	TI *	RR	EE
RR †	Ti	RR *	TE	
Te †	Dg	TF *		Ti †
	Te		Ti †	RR †
Ti	TF	RR	RR †	TE †
Dg	RR		TE †	
TF	EE	Ti †		Ti †
RR	Ti	RR †	Ti *	RR †
Tf	DW	NE †	RR *	TE †
RR	TE		TF *	
		Ti †		Ti †
Ti *	Ti *	RR †	TF	RR †
RR *	RR *	NE †	RR	TE †
TF *	TF *		NE	Ti †
			Ti	RR †
			RR	NE †

Note. For the human coder, the matches are: † = IRE (defined with this sequence of substitutable events: Ti-RR-TE/NE/EE/Te), * = IRF (TI/Ti-RR-TF/BF/Tf), and ‡ = IDE (TI/SI/BI-DD/DG/Dg/DW-TE/Te/SE/Se/EE/ee/BE/Be).

Appendix C
Motif Example

a. Original session codes

Ti RR TF Dg EE Ti DD EE Ti RR EE Ti RR Te Ti Dg TF RR Tf . . . Ti RR TF

b. Cut string into segments of length L = 10 (with overlap)

Ti RR TF Dg EE Ti DD EE Ti RR
Ti RR EE Ti RR Te Ti Dg TF RR
TF RR TF RR Tf RR Ti RR TF Ti
etc

c. Randomly assign starting site and motif selection

Ti RR TF [Dg EE Ti] DD EE Ti RR — Motif 1 at Site 4
Ti RR EE Ti RR [Te Ti Dg] TF RR — Motif 2 at Site 6
TF [RR TF RR] Tf RR Ti RR TF Ti — Motif 1 at Site 2

d. Starting motif's defined

If Motif 1 is made of [Dg EE Ti] and [RR TF RR]

Code	1st site	2nd site	3rd site
Dg	0.50	0	0
EE	0	0.50	0
Ti	0	0	0.50
RR	0.50	0	0.50
TF	0	0.50	0

It is assumed for this example that only the first and third segments are randomly associated with Motif 1 at the start of the algorithm. If Motif 1 is present starting at Sites 4 and 2 respectively, the motif would initially be defined as in the table above with each cells recoding the probability of that site (1^{st}, 2^{nd}, or 3^{rd}) receiving the appropriate code (Dg, EE, Ti, RR, or TF).

***Figure C1*. Motif example.**

Iterative Steps

STEP 1. Select a segment.

Ti RR TF Dg EE Ti DD EE Ti RR

STEP 2. Based on defined motifs determine probabilities for all combinations of motif and starting position.

Motif 1 Site 1: Ti RR TF Probability = 0.01
Motif 1 Site 2: RR TF Dg Probability = 0.01
…
Motif 3 Site 8 EE Ti RR Probability = 0.03

STEP 3. New site and motif selected at random based on probabilities.
So Motif 3 at Site 8 is 3 times as likely as Motif 1 at Site 1.
For demonstration, select Motif 1 starting at Site 6.

Ti RR TF Dg EE [Ti DD EE] Ti RR Motif 1 at Site 6

STEP 4. Redefine motifs and go to STEP 1. (Repeat 20,000 times)

Over time, the definition of each motif settles, and sites where motifs occur are identified. Output is a list of the assignment of motif and starting site for each segment and the definition of the motifs.

Segment	Motif	Start
1	3	4
2	2	6
Last	3	3

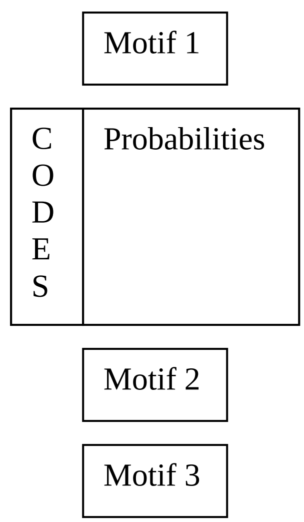

Figure C2. **Motif example, continued.**