

Chapter 6

Collaborative and Differential Utterances, Pivotal Moments, and Polyphony

Stefan Trausan-Matu

Introduction

A deep analysis of collaborative learning sessions should consider several facets. A first aspect is in what degree and how group interactions involved in joint learning provide a scaffold for the individual development. Students have a personal, individual learning trajectory, which interferes with that of the other students when they are learning in a group, like in polyphonic music, where voices have both longitudinal and transversal dimensions.

The dual individual–group perspectives are extremely important for students entering into a process with two cycles, in which they should interact with the others, debate, negotiate meaning in order to construct knowledge, and, meanwhile, internalize it (Stahl, 2006; Vygotsky, 1934/1962). Starting from Bakhtin’s (1981) dialogism ideas we consider dialogue as being essential in both the group and individual cycles: Students enter in dialogues with other students in the first case and with themselves in the second case (for example, the “make problematic” link in Stahl’s cycle of knowledge building (Stahl, 2006)). Moreover, we consider that there is an interaction between external and internal dialogues, for example, external dialogue utterances of one student may have as reaction an internal dialogue utterance at other (or even the same) student, which may be externalized later as an utterance with a loud voice.

Another issue to be considered in collaborative learning is the identification of all types of implied utterances, the role played by words and spoken or written communication but also by other types of communication acts, which may be similar in effect with textual utterances. Natural language is both a means of joint knowledge building and a way for professors towards monitoring the learning process. However, natural language is not the sole way of communication in collaborative learning.

S. Trausan-Matu (✉)
University “Politehnica” of Bucharest, Bucharest, Romania
e-mail: stefan.trausan@cs.pub.ro

In addition to spoken or written language, other means for collaborative knowledge construction may be identified: visual communication, either using diagrams, drawings, images, and objects or body language. All of these may be considered as utterances, in a generalized way, and all may give indicators for learning. However, a big problem is that the set of utterances that may be taken into account is very large, even if we consider only the textual ones. Therefore we should have a means to identify the relevant ones, which are recurrent, have an influence on the learning process, and have an “echo” in the future. As we will later discuss in detail, we call such utterances “voices.”

There were several approaches directed towards the analysis of collaborative learning sessions. Their vast majority considered textual utterances: transcriptions of spoken conversation, logs of instant messenger (chat), forum interventions, and even wikis, for example: CORDTRA (Hmelo-Silver, Chernobilsky, & Masto, 2006), COALA (Dowell & Gladisch, 2007), DIGALO and other tools used in the Argonaut system (Harrer, Hever, & Ziebarth, 2007), and ColAT (Avouris, Fiotakis, Kahrimanis, & Margaritis, 2007). Multimedia utterances were also considered, for example in TATIANA (Dyke, Lund, & Girardot, 2009). Some of these systems use several kinds of argumentation graphs, some of them in the idea of Toulmin (1958), or more elaborated structures like the contingency graphs (Suthers, Dwyer, Medina, & Vatrapu, 2007).

The analysis in existing approaches is usually focused on pairs of utterances: adjacency pairs (Schegloff & Sacks, 1973; Jurafsky & Martin, 2009), transacts (Joshi & Rosé, 2007), or, considering also longer distance connections, uptakes (Suthers et al., 2007). We consider that also another, more global unit of interaction than that of pairs of utterances should be considered: threads of utterances interanimating in a polyphonic framework (Trausan-Matu, Stahl, & Sarmiento, 2007). For example, even the significance of a pair of utterances may be totally different if they are singular or if the second utterance in the pair appears after a thread of repetitions of the first utterance. Moreover, repetitions of utterances, either singular or in pairs, may generate a rhythm.

Utterances which are influential become “voices” that means threads having a duration and/or echoes. In our vision, an utterance, in a generalized sense (and consequently a potential voice), may be a word, a sentence, a paragraph, a paper, a book, a turn in a conversation, a figure, a gesture, etc. Utterances may be not only individual, but they may also be generated by a group (for example, all students move their chairs as a chorus at the beginning of the origami fractions session).

Learning, either individual or collaborative, has duration (a longitudinal dimension in time) and can occur at different rhythms (even in the same session) of dialogue and in different settings. Consequently, a derived problem is what types of space–time situations or chronotopes (Bakhtin, 1981; Ligorio & Ritella, 2010) may be identified in the analyzed data, and in what degree they are well suited for achieving a good collaboration.¹ For example, in the beginning of collaborative problem solving a chronotope with few verbal utterances may be detected, in which

¹In our approach, we consider a chronotope as “a genre of movement or pacing in the space that participants adopt over the temporal duration of an activity” (Ligorio & Ritella, 2010).

students explore the problem. When they reach to build collaboratively a solution, another chronotope may be detected, which may be called also a region of good collaboration (Banica, Trausan-Matu, & Rebedea, 2011), in which threads of verbal utterances occur in a rapid rhythm.

Changes in learning rhythm are the starting point for passing from one chronotope to another (Ligorio & Ritella, 2010) and may be considered pivotal moments in the learning session. Changes of rhythm are often associated with the presence of special utterances, for example, collaborative or differential (Trausan-Matu et al., 2007), which may be therefore considered as cues for detecting pivotal moments. Collaborative utterances, even if they sometimes don't mark a change of rhythm (a passage from a chronotope to another) are also candidates for pivotal moments because they are not frequent situations and they display moments in which the group behaves like a whole; it really collaborates, which is a desideratum in computer-supported collaborative learning.

A good professor is able to orchestrate utterances as voices: he/she gives texts to students to read, speaks, uses images and gestures, and even analyzes and directs the class's acts (or utterances as a group) in order to build a coherent thread of ideas. This process is similar to music not only by the existence of a polyphony of voices but also through the created rhythms.

The identification of the types of chronotopes, of collaborative moments, and of pivotal moments in a learning session are very important for a teacher in order to manage students' activity. A model which can provide a unifying view on the above facets is multivocality and polyphony (Trausan-Matu et al., 2007), which will be used in this chapter for analyzing the origami fractions data set. This model may also be used to implement semiautomatic analysis tools, which provide facilities for the visualization of voices and their interanimation and potential pivotal moments (Trausan-Matu & Rebedea, 2009; Chiru & Trausan-Matu, 2012).

The Five Dimensions Characterizing the Approach

The method of analysis of collaborative learning used in this chapter is based on considering small-group interactions from the perspectives of dialogism and polyphony (Bakhtin, 1981, 1984; Trausan-Matu et al., 2007), repetition and rhythm as an involvement provider (Tannen, 1989), interanimation (Wegerif, 2005; Trausan-Matu et al., 2007), conversation analysis (Sacks, 1962/1995)—collaborative utterances and adjacency pairs), and collaborative moments (Stahl, 2006). The five dimensions on which our approach may be understood are the following:

Assumptions Underlying the Analysis

Theoretical assumptions. Knowledge may be constructed in small groups (Stahl, 2006, 2009). In this process, interplays take place between the group discourse and the understanding of the participants as individuals (Stahl, 2006).

Small-group conversations for problem solving and collaborative learning often take the form of multi-threaded discourse that follows polyphonic patterns (Trausan-Matu et al., 2007). Both group discourse and individual thinking are characterized by dialogism and multivocality (Bakhtin, 1984; Trausan-Matu et al., 2007).

Methodological assumptions. Interanimation patterns (Trausan-Matu et al., 2007) may be detected in interactions, and they offer a glimpse on the collaborative learning processes of the group. Conversation analysis and ethnomethodology (Garfinkel, 1967) may be used for providing cues for detecting interanimation and collaboration (by identification of associated member methods). Integrating natural language processing (NLP) techniques (for the automatic identification of adjacency pairs, repetition, and discourse threads) with polyphony identification, social network analysis, and graphical visualizations may provide a way for analyzing the contributions of each participant and their interanimation.

Purpose of Analysis

A main purpose of analysis from the point of view of this chapter's approach is the recognition of interanimation patterns among voices (in particular considering participants and discussion threads) and, as a result, the inference of pivotal moments mentioned earlier and regions of good collaboration. Related purposes are the identification of collaborative and differential utterances, of adjacency pairs, of voices (discourse threads) and their interactions, and of the semantic and pragmatic content of the utterances. Eventually, starting from the above data, an evaluation of the contribution of each participant to the learning process may be also derived.

Units of Interaction

The most important units of interaction in our approach are voices, in a generalized sense, which means, from another perspective, discourse threads viewed in a polyphonic weaving. However, as units of interaction are also considered pairs of utterances. We remind that utterances, in a generalized sense may be: words, sentences, gestures, and images. All these may be seen also as units of action.

Representations of Data and Analytic Interpretations

Transcriptions of textual utterances are codified using a complex XML schema in order to be available for an automatic analysis. Graphical representations of some types of voices and their interanimation are generated automatically. Graphical representation of the evolution of the contribution of each participant may also be represented.

Analytic Manipulations

There are two main analysis directions. The first of them is the analysis of discourse for identifying voices, repetitions of generalized utterances (as defined above) in order to construct threads, and their interactions. This objective includes the analysis of speech acts, adjacency pairs, collaborative and differential utterances, co-references, argumentation chains, contrapuntal/polyphonic structure, etc. and (if available) nonverbal communication and individual/group body language. NLP tools are used as a support of the analysis. The second direction is the analysis of the social network of user links between their utterances.

The Polyphonic Model and the Interanimation Patterns

Polyphony is an example of a joint achievement of several independent participants acting sequentially (singing in music or emitting utterances in dialogues) starting from a common theme and meanwhile keeping coherence among them. It originated as a concept and practice in music, and it can be extrapolated to texts, as Bakhtin (1984) emphasized and even, in our opinion, to spoken and nonverbal artifacts. Polyphony may occur in musical pieces with more than one melodic line (or voice) at a time, in contrast with monophony, where a single voice (part) is present. Polyphony differs also from homophony because even if in both cases multiple voices are present, in the former they have a high degree of independence. However, even if they are independent, in order to achieve polyphony, the voices obey some implicit constraints, some so-called counterpoint rules, for example, in order to achieve a joint harmonic, pleasant musical piece. Polyphony may be seen as a model of group interaction and creativity, in which independent individuals (voices, in a metaphorical sense) achieve a joint activity during a period of time.

We consider a voice, in a generalized way, beyond the acoustic sense, as a distinctive presence in a group, influencing the other voices. An utterance or a sequence of utterances become a voice if they have a longitudinal dimension, they last, they have an echo in time, and they may be perceived as a coherent thread. Meanwhile, to have a distinctive presence, a voice should have a transversal dimension, opposing but also keeping coherence with the other voices.

One important feature of the polyphonic model is that if we consider the generalized perspective of a voice, it may be applied for an integrated analysis of different types of media for communication. Even if it was conceived by integrating ideas from music and text, it may be applied to analyzing video images, as will be the case in the analysis of the origami fractions data set presented in this chapter. A voice in our polyphonic model may be a spoken utterance, a written utterance on the blackboard, but also nonverbal utterances like a gaze, a movement, or the acts of the teacher, a student, or a group of students.

The polyphonic model of group interaction in collaborative learning considers that in a conversation different longitudinal threads (or “voices”) appear, composed of utterances and their echoes, each of them having independence but achieving a joint (a consonant) discourse (Trausan-Matu et al., 2007; Trausan-Matu & Rebedea, 2009). However, the interaction in a group inherently involves the solving of the dissonances appearing between voices. Therefore, as also Bakhtin noticed for texts, in general (Bakhtin, 1981), participants face both centrifugal (divergent, towards difference) and centripetal (convergent, towards unity) forces, along two directions: longitudinal and transversal, following constraints that are similar to the music counterpoint rules (Trausan-Matu et al., 2007). These forces have an important effect: they obligate the participants to perceive dissonances that put their utterances under question (they make them “problematic” in the personal cycle of the knowledge building (Stahl, 2006)), and they generate an interanimation phenomenon. The polyphonic analysis tries to identify interanimation patterns along the two dimensions while corresponding to the two types of forces.

The polyphonic analysis of a joint activity like those specific to collaborative learning combines the individual and group perspectives. Similarly to the case of participants in an improvising jazz quartet, each learner is listening to (and sometimes also looking at) the others and is also playing in the same time, achieving a joint musical piece. It is very important to consider the group as a whole—not just individual developments or dyadic interactions within the group. The joint achievement of the group, be it music or spoken or written dialogue, is constrained by the centripetal and centrifugal forces towards convergence/divergence, and it may be seen as a creative or a “thinking device” (Wegerif, 2005). The presence of centrifugal and centripetal forces may be discovered by the identification of interanimation patterns among participants’ utterances.

Interanimation patterns may be classified in unity-pursuing patterns, characterized by a trend towards continuity and achieving coherence in the interaction and differential interanimation patterns (Trausan-Matu et al., 2007). They may be identified, for example, in transcriptions or chat logs using conversation analysis (CA—Sacks, 1962/1995) or NLP, and they may be the starting point for analyzing the degree of collaboration and personal contributions (Trausan-Matu & Rebedea, 2010). Interanimation patterns occur also in face-to-face interaction, including non-verbal behavior, as will be discussed later in this chapter.

A very important case of unity patterns is the cumulative talk (Mercer, 2000) or, in Sacks’ words, collaborative utterances (Sacks, 1962/1995). This type of convergent interaction is characterized by the fact that two or more participants spontaneously build together a sentence, as if they were a single person. Two examples are found in Sacks (1962/1995):

Joe	(Coughs) We were in an automobile discussion,
Henry	discussing the psychological motives for
Mel drag	racing on the streets

and Trausan-Matu et al. (2007):

ModeratorSf	Could you guys tell templar what's going on?
Mathpudding	We're experimenting with circles
Mathman	and finding as many possible relations as we can

This kind of pattern occurs also in the data set, at several points, for example, at utterances 457–459, one of the pivotal moments is as follows:

34:40	457	N	Although the production methods differ [starts quietly],
	458	T	Yes.
	459	G	The shape is the same.
		K	The shape is the same.
		N	The shape is the same.

Collaborative utterances appear in several places in the origami fractions data set. They are rare, and they are generally related to pivotal moments (which might be related to what Stahl (2006) calls “collaboration moments”) in which the group displays cohesion and sometimes understanding. Collaborative utterances may also be nonverbal, in body language, like the fact that everybody (excepting Y, see next section for details) moves the chairs as if it were a choreography, at moment 0:25² in the video.

If collaborative utterances might be considered examples of consonances in the polyphonic metaphor, differential patterns may be viewed as examples of dissonances, of something felt as unfinalized or wrong. They have a very important role in triggering further utterances of other participants as a result of incompleteness perception.

A differential pattern example is (taken from Stahl, 2006 and commented in Trausan-Matu & Rebedea, 2009) the following:

1:21:53	Teacher: And you don't have anything like that there?
1:21:56	Steven: I don't think so
1:21:57	Jamie: Not with the same engine
1:21:58	Steven: \lceil No Jamie: \lfloor Not with the same
1:21:59	Teacher: With the same engine ... but with a different (0.1) ... nose cone?= 1:22:01
1:22:01	Chuck: \lceil =The same= Jamie: \lfloor =Yeah,
1:22:02	Chuck: These are both (0.8) the same thing
1:22:04	Teacher: Aw \lceil right
1:22:05	Brent: \lfloor This one's different

Remark that this differential pattern occurs after a series of repetitions of “the same” which becomes a thread or, in other perspective, a voice inducing a dissonance needing a resolution.

²The indicated times are those from the video file, not those in the transcription or subtitling, which are slightly different.

Differential patterns are also essential for the identification of pivotal moments in the origami fractions data set. An important fact to remark is that the below examples of differential patterns are connected to the collaborative utterance (473–474) marking the first pivotal moment. Note that almost the same words occur in the above and below excerpts from two different corpora:

35:58	469	T	What do you think of N's two solutions? [Places N's two solutions on the teacher's desk.]
	470	Y	[Moves towards the teacher's desk by further raising his hip.]
	471	Anonymous	[Whispers] The shapes differ.
36:14	472	Y	Differ [with clear voice]
	473	Y	though areas are equal [with low voice].
	474	G	The areas are the same,
	475	T	Yes.
36:20	476	G	but the shapes and production methods differ.

Differential patterns may occur also (as in the case of collaborative utterances) in body language, as it will be discussed in a section below.

The Analysis of the Origami Fractions Data Set

Many unity and differential interanimation patterns of different kinds, on different dimensions (verbal and nonverbal), may be identified in the origami fractions data set, occurring among different types of voices: participants' spoken utterances, body language utterances, solutions, opinions, threads of repeated words, etc. Some of the interanimation patterns are unprompted (for example, the collaborative utterances) and some are induced by the teacher (for example, threads of repeated differential patterns aimed at inducing the answer to the problem). In general, teachers should know how to handle voices and interanimation patterns. They have to be able to detect collaborative utterances that may be a sign of moments of collaboration. The repeating of difference patterns may induce understanding. Different kinds of additional voices like images or drawings may be used for inducing interanimation.

Pivotal moments in our perspective are generally associated to the presence of collaborative or differential utterances, which occur many times as a result of threads' (voices') interaction. As we mentioned previously, pivotal moments (and collaborative and differential utterances) also coincide sometimes with changes in the learning rhythm, marking the passing from one chronotope to another (Ligorio & Ritella, 2010).

I discovered some of the instances of interanimation patterns on a later, more thorough analysis, after seeing that Chiu's analysis (Chap. 7, [this volume](#)) contained more pivotal moments than mine. Moreover, his discussion on micro-creativity enforced me the unifying view of CSCL and computer-supported group creativity under the polyphonic model.

We analyze the origami fractions data set, according to several dimensions which can be considered intertwined, following the polyphonic model. The dimensions we consider in the following sections are spoken dialogue, body language, visual dimension,

internal dialogue (at an intramental level), and echoes. In each of these dimensions, several voices, in a metaphorical way, and their polyphonic interactions may be detected.

Spoken Dialogue

The first and probably most important dimension consists of individual and collaborative utterances in the spoken dialogue. This dimension may be investigated by CA (Sacks, 1962/1995), discourse analysis (Tannen, 1989), interanimation (Trausan-Matu et al., 2007; Trausan-Matu & Rebedea, 2009), and NLP methods (Trausan-Matu & Rebedea, 2010).

As mentioned in the previous section, collaborative and differential patterns may be detected in the transcribed data of the origami fractions session. A very “dense” segment, with several collaborative and differential utterances, is between utterances 469 and 482. The segment starts with the first pivotal moment labeled by Shirouzu’s analysis (Shirouzu, Chap. 5, [this volume](#)) which is also the fifth of Chiu (472–474) (Chiu, Chap. 7, [this volume](#)). We also identified this segment as a pivotal moment within our polyphonic perspective due to both differential interanimation patterns (at 471 and 476) and collaborative utterances (473–474, 476, and 481–482).

35:58	469	T	What do you think of N’s two solutions? [Places N’s two solutions on the teacher’s desk.]
	470	Y	[Moves towards the teacher’s desk by further raising his hip.]
	471	Anonymous	[Whispers] The shapes differ.
36:14	472	Y	Differ [with clear voice]
	473	Y	though areas are equal [with low voice].
	474	G	The areas are the same,
	475	T	Yes.
36:20	476	G	but the shapes and production methods differ.
		K	The shape and production method differ.
		N	The shape and production method differ.
		Anonymous	The shape and production method differ.
	477	T	The areas are the same.
	478	T	Because the areas are the same,
	479	T	this is the last comparison [N’s first solution and K’s one].
		Y	[Leans over the desk.]
36:52	480	T	What do you think of these?
37:04	481	G	Although shapes are the same,
		K	Although shapes are the same,
		N	Although the shapes are the same,
		O	Although the shapes are the same,
		Y	Although the shapes are the same,
	482	G	the production methods differ.
		K	the production methods differ.
		N	the production methods differ.
		O	the production methods differ.
		Y	[Quickly goes back to his seat.]

The second pivotal moment identified by Shirouzu (Shirouzu, Chap. 5, [this volume](#)) corresponds also to collaborative utterances 502–503:

38:13	495	T	What among these is constant?
38:14	496	Anonymous	[Whispers] Area.
	497	Anonymous	[All together] Area.
	498	T	Area.
	499	T	The areas are the same.
38:18	500	T	How large is the area?
38:20	501	Anonymous	[Whispers] of 2 (halves), 2 ...
38:24	502	Y	1/2 [in low voice]
	503	Y	[Slowly] of the whole.
		T	[Following Y] 1/2 of the whole
	504	G	Ah.
		K	Ah. [Moves right hand.]
		N	Ah. [Nods.]

Other collaborative utterances occur in several places of the data set. For example, at utterances 8–10, a first verbal joint, collaborative utterance marks the beginning of the problem solving:

0:00	1	T	Here we have a piece of origami paper, a pencil, and a pair of scissors.
	2	T	What I want you to do is ...
	3	T	to use these to make three-fourths of two-thirds of this origami paper.
0:27	4	T	Can anybody do that?
0:30	5	N	Can I?
	6	T	Oh, you need this? [Handing a piece of origami paper to N.]
	7	N	[Starts to fold the paper into a rectangle of one-third of the total area.]
	8	F	Of two-thirds ...
	9	G	Of two-thirds ...
	10	K	Three-fourths.

Differential patterns may also be considered for detecting pivotal moments, as mentioned above. They occur sometimes after a series of repetitions (as in the collaborative moment in the solving of the rocket nose problem in Stahl (2006)) and/or together a collaborative utterance, like at utterances 471–476 in the Shirouzu (Shirouzu, Chap. 4, [this volume](#)) data set.

Body Language

The second dimension of analysis that we consider is body language, which may contain individual or collective utterances. An example is the moment (at 0:25) when, after the teacher appears in front of the students, all of them move their chairs forward,

excepting Y. Such a movement seemed like a collective spontaneous sign of their entering into the lesson space. Ethnomethodology may be used for analyzing such member methods (Garfinkel, 1967). The Y student's body language is in many moments independent, behaving like a distinct voice, on a differential pattern. He stays almost immobile for the majority of the first 30 min. An important moment is at about 36:37, when for several tens of seconds Y stands up, moves towards the table, and looks transversally. This is important because this moment coincides with another crucial moment (pivotal moment 1), when student Y has a very important contribution.

Collective body language is also displayed by some students (N, F, and K) when they avoid answering the teacher's question:

561	T	I would now like to ask each of you what you did,
562	T	because all of you have solved it.

They do this by putting their hands over their eyes, putting their heads on the table, or looking elsewhere (all different "methods" of avoiding an answer that are very well known by professors).

The reaction of the students may be viewed also as a voice saying "we don't want to answer any more," and it could even be considered as a pivotal moment in the lesson, possibly indicating students' fatigue and thus the beginning of another chronotope. As a consequence, the teacher does not insist and answers himself.

The Visual Dimension

What participants see is a third dimension of analysis. Visual data on the blackboard, what other participants do, and even others' body language are "voices" that may generate reactions that may be sparks triggering interanimation patterns. The actions of the teacher that writes on the table and displays the solutions may be seen as voices that are supposed to trigger students' internal reasoning and responses.

Shirouzu (Chap. 4, [this volume](#)) mentions that the origami fractions experiment had two phases. In the first 30 min, children were instructed to solve the problem of "obtaining $3/4$ of $2/3$ of colored paper (origami paper)" using provided colored paper and scissors. This process occurred mainly individually, although a joint component is present because they could look at each other and compare their utterances (including origami folding and cutting acts), seeing who solved the problem and how they went about it. The visual dimension was enforced by the teacher when displaying the solutions on the table.

Intramental Dimension

In the beginning part of the data set some students are folding and cutting origami (G and N) and some are watching (Y, K, and O). After others started to individually solve the problem by cutting and folding origami, Y proposed a solution (at 13.07) totally

different from the others. He also has the major contributions at pivotal moments 2 and 3, and after 5 months he was the student having the best description of the origami session findings. One explanation of his achievements in spite of his predominant less active participation might be that he is probably rather an intramental than an intermental reasoner, a lurker positioning himself on differential positions (as discussed in the body language section) even regarding his own verbal utterances (doubting at utterance 519 that what he said before was right: “Is this wrong?”). At least in this data set, his actions show that he prefers to look and afterward to act. In polyphony terms, he prefers to develop a counterpoint while internalizing others’ voices and to have inner dialogues rather than entering into polyphony with the others. Even the fact that he does not stay at the table for the majority of time is perhaps an argument for our idea. Based on what we can observe in the origami fractions data set we could say that Y is a divergent thinker.

A similar assertion may be said partially about K, who inversed the order of fractions (2/3 of 3/4 instead of 3/4 of 2/3). We may remark also that K, from the beginning, has a different position:

8	F	Of two-thirds ...
9	G	Of two-thirds ...
10	K	Three-fourths.

It is interesting to note that even if K seems to inverse the order, she also had one of the best rememberings of the idea of the session after 5 months. In another interpretation of utterance 10, K might be completing the previous utterances according to teacher’s specification.

The ideas of inner speech and dialogue have an important role in the writings of Vygotsky (1934/1962) and Bakhtin (Voloshinov, 1929/1973). For example, Bakhtin says: “There are no ontological differences between inner and outer speech” (Clark & Holquist, 1984).

Stahl’s personal understanding cycle contains also inner acts: “We may be able to repair our understanding by explicating the implications of that understanding and resolving conflicts or filling in gaps—by reinterpreting our meaning structures—to arrive at a new comprehension” (Stahl, 2006). He considers that what happens at the individual mind level is socially determined: “The process of interpretation that seems to be carried out at the level of the individual mind is already an essentially social process” (Stahl, 2006).

Some neurology researchers are also supporting the idea of inner speech, following the ideas of the Russian school initiated by Vygotsky and continued by Luria (DeBleser & Marshall, 2005). Neural correlates of inner speech are also mentioned (Jones & Fernyhough, 2007). I searched such evidences after the “polyphonic interanimation” of my, Shirouzu’s, and Lund’s ideas and opinions related to the intramental dimension, exchanged during our interactions around the data set.

Thinking—the intramental activity—is, in our vision dialogical, implying inner speech which, similarly to the outer speech, is composed of inner utterances. If we consider that there is such a dimension, at least two types of students’ thinking may

be supposed to be present in the origami data set. The first one is that occurring when they are individually trying to obtain the solution by folding origami following the verbalized goal specified by the professor: “to make $3/4$ of $2/3$ of colored (origami) paper.” In support of this idea we remark that they have to achieve at least two sequential steps (obtain $2/3$ and $3/4$) and therefore to propose actions, to remember them, and to validate their correctness, all of these made without loud voice. We may say that they have to emit inner utterances like “I fold ...” or “I cut ...” Such utterances might not be linguistic; they might be generalized utterances and mental imagery of the folding, cutting, and comparing acts.

In order to solve the problem students should also emit inner utterances in a kind of inner dialogue with themselves, containing sentences as “the (partial) result is good/wrong” or adjacency pairs like question–answer.

A second type of utterance at the intramental dimension is generated by looking at others’ solutions, at the teacher’s writing, and at the display of solutions on the blackboard. Hearing teacher’s and others’ utterances probably also generates inner utterances (for example, “my solution is the same as ...” or “my solution is different from ...”), and interanimation patterns may occur (for example, we can consider adjacency pairs (Schegloff & Sacks, 1973) between external and internal utterances, which might also be uptakes (Suthers et al., 2007)). Other types of thinking may be identified, for example, to prepare an answer to teacher’s questions and even the attempts to avoid an answer (N, F, and K after the teacher’s utterances 561 and 562).

Echoes

The fifth analysis dimension in our approach is the long-term effect, the long-term echo of the voices, spoken, inner, or of another kind, which were present in the lesson. This dimension is very important because it is, in fact, the main goal of the teaching session. The analysis made after 5 months shows that either students forgot or did not initially understand the conclusion of the lesson. After 5 months, only Y, who proposed the solution, and K remembered the final conclusion (Y: “The $2/3 \times 3/4$ made $1/2$ and we were taught why it resulted in $1/2$ ”; K: “We thought why $2/3 \times 3/4$ equals $1/2$ ”). An answer to the question “why was there a difference between Child Y and G, in spite of G’s convergent moves to Y at pivotal moments 1 and 2” may be given starting from the idea of collaborative utterances. G acted as a member of a group, participating in collaborative utterances, but she didn’t internalize the utterance; she only acted as a “mirror” (see Tannen, 1989).

Tools for Helping the Polyphonic Analysis

In the analysis presented in this chapter, the detection of pivotal moments was based primarily on a manual analysis towards the identification of interanimation patterns and the identification of changes in rhythm (passing from a chronotope to another),

which sometimes co-occur. The automatic detection of voices, of the instances of interanimation patterns, and of polyphony would be extremely useful, but it is extremely difficult, even if only for textual utterances. An easier task is to assist a human analyst by trying to identify specific behaviors that may indicate the possible presence of voices, interanimation patterns, and changes of rhythm. For example, it is easy to detect repeating words or phrases which may signal a thread, a voice. Moreover, discourse markers, cue phrases, and particular speech acts may be used for detecting differential patterns.

The semiautomatic content-based analysis system PolyCAFe (Polyphonic Conversation Analysis and Feedback generation) proved helpful for the analysis of collaborative learning sessions using instant messenger chats (Trausan-Matu & Rebedea, 2010). This system is based on the polyphonic model (Bakhtin, 1981, 1984; Trausan-Matu et al., 2007) and assists human analysts in the detection and the visualization of the presence of voices, interanimation patterns, participation, contribution, semantic content, and collaboration in conversations (Trausan-Matu & Rebedea, 2010; Rebedea, Dascalu, Trausan-Matu, Armit, & Chiru, 2011; see also <http://www.ltfll-project.org/index.php/polycafe.html>). The system uses techniques from NLP and social network analysis (Dascalu, Rebedea, & Trausan-Matu, 2010; Rebedea et al., 2011).

PolyCAFe is a module developed in the EU FP7-IST project “Language Technologies for Lifelong Learning” (LTfLL, see <http://www.ltfll-project.org>), and it provides textual feedback and interactive graphic visualization of instant messenger chats, transcribed conversations, forums, or other collaborative activities. The system offers (among other services) facilities for the identification of adjacency pairs, for identification of the most frequent used concepts, and for the visualization of threads (voices) and their interactions.

PolyCAFe was used for the analysis of the transcription of the discussions in the origami fractions data set. For this purpose, the transcription³ was encoded into a specific XML schema, processed, and analyzed with the graphical facilities. Figure 6.1 shows the graphical visualization of the threading generated by the words “different,” “same,” “solution,” and “number” in the origami fractions data set, which may be used for the identification of some interanimation patterns (each participant’s utterances are small rectangles on a horizontal line, time flowing from left to right; the threads of appearance of concepts (words) are shown with distinct colors; the ruler shows the number of utterances). For example, after several rhythmical repetitions of the word “same,” a joint appearance of “same” and “different” occurs after utterance 480.

³Of course such a tool loses many useful details of the face-to-face origami data set because it is not able to identify nonverbal utterances. However, as it will be seen below, it still provides some useful representation for analyzing rhythm and pivotal moments.

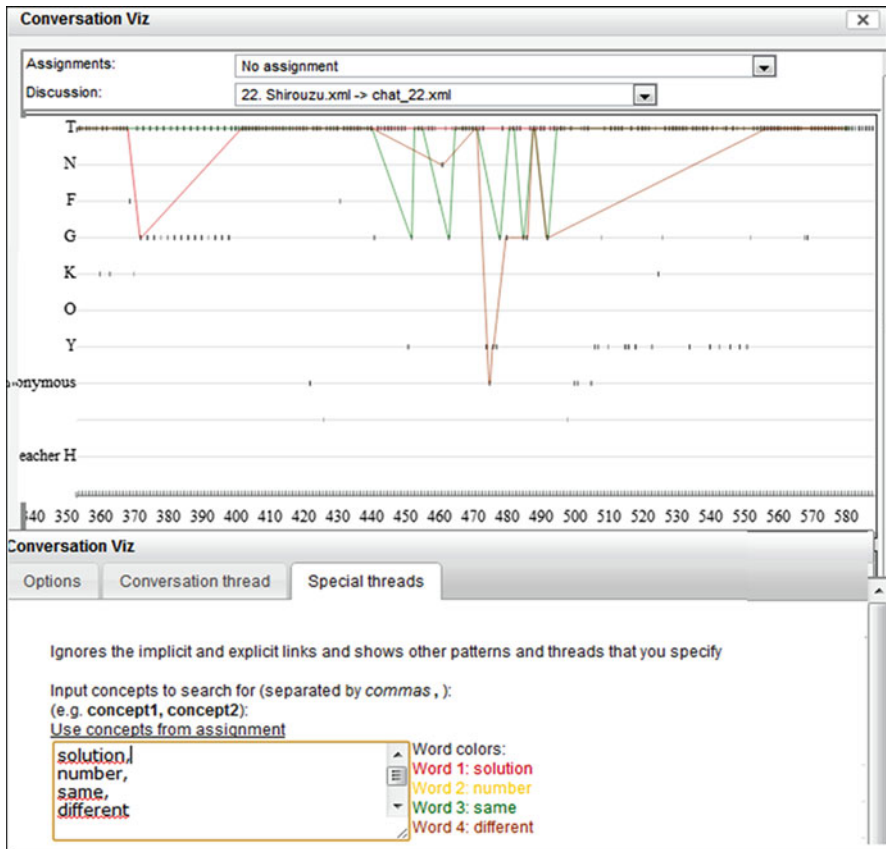


Fig. 6.1 The PolyCAFe visualization of a fragment of the conversation in the origami fractions data set

Conclusions

Pivotal moments in the approach presented in this chapter are related to collaborative moments (collaborative utterances), to other interanimation patterns (for example, differential utterances), and sometimes to changes in the rhythm (chronotope) of interacting voices. The analysis presented showed that the detection of pivotal moments in conversations may start from the identification of two types of interanimation patterns: collaborative and differential utterances and their succession.

In the origami fractions data set the pivotal moments that can be detected by the polyphonic approach are in the first minute (collaborative utterances both spoken 8–10 and collective body language), at the first pivotal moment of Shirouzu

(collaborative and differential utterances), at the second pivotal moment of Shirouzu (collaborative utterances), and at the third pivotal moment of Shirouzu (the 548 collaborative utterance). Another possible pivotal moment is at utterances 561–562 (body language).

An important achievement of the analysis of the origami data set with the polyphonic model was the natural extension of its usage beyond textual utterances. Voices and interanimation patterns were identified also between verbal and nonverbal utterances. The concept of generalized utterances was introduced in order to include visual utterances, body language, and group utterances. Moreover, the existence of the intramental dimension that includes inner dialogue and inner utterances was asserted because it may explain some observed facts in the data set. The assertion of this dimension is also based on the ideas of inner speech (Vygotsky, 1934/1962; DeBleser & Marshall, 2005; Jones & Fernyhough, 2007), inner dialogue (Voloshinov, 1929/1973), and personal understanding cycle (Stahl, 2006). However, further investigations and evidence are needed for this latter dimension.

Acknowledgements I would like to thank Kris Lund and Dan Suthers for their very useful comments and suggestions. I also want to thank Hajime Shirouzu for the very interesting data set; Hajime and Ming Ming Chiu for the very interesting discussions on the analysis of the learning sessions; and Costin Chiru, Traian Rebedea, and Mihai Dascalu for their help on the automated support development. The work in this chapter was supported partially by the EU FP7 projects LTFLL and ERRIC.

References

- Avouris, N., Fiotakis, G., Kahrimanis, G., & Margaritis, M. (2007). Beyond Logging of Fingertip Actions: Analysis of Collaborative Learning Using Multiple Sources of Data. *Journal of Interactive Learning Research*, 18, 231–50.
- Bakhtin, M. (1981). *The dialogic imagination: Four essays*. London: University of Texas Press.
- Bakhtin, M. (1984). *Problems of Dostoevsky's poetics (C. Emerson, Trans.)*. Minneapolis, MN: University of Minnesota Press.
- Banica, D., Trausan-Matu, S., & Rebedea, T. (2011). Detecting collaboration regions in a chat session. In H. Spada, G. Stahl, N. Miyake, N. Law (Eds.), *CSCL 2011 Conference Proceedings* (Vol. 2, pp. 771–775). Hong Kong.
- Chiru, C., & Trausan-Matu, S. (2012) Identification and Classification of the Most Important Moments from Students' Collaborative Discourses. In *Proceedings of ITS 2012*. Springer
- Chiu, M. M. (this volume). Social metacognition, micro-creativity and justifications: Statistical discourse analysis of a mathematics classroom conversation. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs & N. Law (Eds.), *Productive Multivocality in the Analysis of Group Interactions*, Chapter 7. New York: Springer.
- Clark, K., & Holquist, J. M. (1984). *Mikhail Bakhtin*. Cambridge, MA: Harvard University Press.
- Dascalu, M., Rebedea, T., & Trausan-Matu, S. (2010). A Deep Insight in Chat Analysis: Collaboration, Evolution and Evaluation, Summarization and Search. *AIMSA 2010, Lecture Notes on Artificial Intelligence* (Vol. 6304, pp. 191–200). Springer.
- DeBleser, R., & Marshall, J. C. (2005). Egon Weigl and the Concept of Inner Speech. *Cortex*, 41, 249–57.
- Dowell, J., & Gladisch, T. (2007). Design of argument diagramming for case-based group learning. *ACM International Conference Proceeding Series* (Vol. 250, 99–105).

- Dyke, G., Lund, K., & Girardot, J.-J. (2009). Tatiana: an environment to support the CSCL analysis process in O'Malley, C., Suthers D., Reimann P., Dimitracopoulou A. (Eds.) *Computer Supported Collaborative Learning Practices - CSCL2009 Conference Proceedings*, Rhodes, Greece, 58–67.
- Garfinkel, H. (1967). *Studies in ethnomethodology*. Englewood Cliffs, NJ: Prentice-Hall.
- Harrer, A., Hever, R., & Ziebarth, S. (2007). Empowering researchers to detect interaction patterns in e-collaboration. *Frontiers in Artificial Intelligence and Applications*, 158, 503–10.
- Hmelo-Silver, C.E., Chernobilsky, E., & Mastro, O. (2006). *Representations for Analyzing Tool-mediated Collaborative Learning. Proceedings of ICLS 2006*. New Brunswick.
- Jones, S. R., & Fernyhough, C. (2007). Neural correlates of inner speech and auditory verbal hallucinations: A critical review and theoretical integration. *Clinical Psychology Review*, 27, 140–54.
- Joshi, M., & Rosé, C.P. (2007). Using transactivity in conversation summarization in educational dialog. In *Proceedings of the SLATE Workshop on Speech and Language Technology in Education*
- Jurafsky, D., & Martin, J.H. (2009) *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition* (2nd edn.), London: Pearson Prentice Hall
- Ligorio, M. B., & Ritella, G. (2010). The collaborative construction of chronotopes during computer supported collaborative professional tasks. *International Journal on Computer-Supported Collaborative Learning*, 5(4), 433–45. Springer.
- Mercer, N. (2000). *Words and minds. How we use language to think together*. London: Routledge.
- Rebedea, T., Dascalu, M., Trausan-Matu, S., Armit, G., & Chiru, C. (2011). Automatic Assessment of Collaborative Chat Conversations with PolyCAFe. In C. Delgado Kloos et al. (Eds.), *EC-TEL 2011* (LNCS, Vol. 6964, pp. 299–312). Berlin: Springer.
- Sacks, H. (1962/1995). *Lectures on conversation*. Oxford: Blackwell.
- Schegloff, E. A., & Sacks, H. (1973). Opening up closings. *Semiotica*, 8, 289–327.
- Shirouzu, H. (this volume). Learning fractions through folding in an elementary face-to-face classroom. In D. D. Suthers, K. Lund, C. P. Rosé, C. Teplovs & N. Law (Eds.), *Productive Multivocality in the Analysis of Group Interactions*, Chapter 5. New York: Springer.
- Stahl, G. (2006). *Group cognition: Computer support for building collaborative knowledge*. Cambridge, MA: MIT Press.
- Stahl, G. (Ed.). (2009). *Studying virtual math teams*. Boston, MA: Springer.
- Suthers, D. D., Dwyer, N., Medina, R., & Vatrpu, R. (2007). A framework for eclectic analysis of collaborative interaction. In C. Chinn, G. Erkens & S. Puntambekar (Eds.), *The Computer Supported Collaborative Learning (CSCL) Conference 2007* (pp. 694–703). New Brunswick: International Society of the Learning Sciences.
- Tannen, D. (1989). *Talking voices: Repetition, dialogue, and imagery in conversational discourse*. Cambridge: Cambridge University Press.
- Toulmin, S. (1958). *The uses of argument*. Cambridge: University Press.
- Trausan-Matu, S., & Rebedea, T. (2009). Polyphonic Inter-Animation of Voices in VMT. In G. Stahl (Ed.), *Studying Virtual Math Teams* (pp. 451–473). Boston, MA: Springer.
- Trausan-Matu, S., & Rebedea, T. (2010). A Polyphonic Model and System for Inter-animation Analysis in Chat Conversations with Multiple Participants. In A. Gelbukh (Ed.), *CICLING 2010* (LNCS, Vol. 6008, pp. 354–363). Heidelberg: Springer.
- Trausan-Matu, S., Stahl, G., & Sarmiento, J. (2007). Supporting polyphonic collaborative learning. *E-service Journal* (Indiana University Press), 6(1), 58–74.
- Voloshinov, V. N. (1929/1973). *Marxism and the philosophy of language*. New York, NY: Seminar Press.
- Vygotsky, L. S. (1934/1962). *Thought and language*. Cambridge, MA: MIT Press.
- Wegerif, R. (2005). Towards a dialogic understanding of the relationship between CSCL and teaching thinking skills. In *Proceedings of the 2005 conference on Computer support for collaborative learning: learning 2005: the next 10 years!* (CSCL '05) (pp. 707–16). International Society of the Learning Sciences.