Article

Theoretical and Methodological Perspectives on Designing Video Studies of Interaction

Anna-Lena Rostvall
Stockholm Institute of Education
Stockholm, Sweden

Tore West
Stockholm Institute of Education
Stockholm, Sweden

Abstract

In this article the authors discuss the theoretical basis for the methodological decisions made during the course of a Swedish research project on interaction and learning. The purpose is to discuss how different theories are applied at separate levels of the study. The study is structured on three levels, with separate sets of research questions and theoretical concepts. The levels reflect a close-up description, a systematic analysis, and an interpretation of how teachers and students act and interact. The data consist of 12 hours of video-recorded and transcribed music lessons from high school and college. Through a multidisciplinary theoretical framework, the general understanding of teaching and learning in terms of interaction can be widened. The authors also present a software tool developed to facilitate the processes of transcription and analysis of the video data.

Keywords: video analysis, interaction, education, multimodality, qualitative data analysis (QDA)

Introduction

Teacher-student interaction and its implications for musical learning are brought into focus in a Swedish study of instrumental teaching described in this article. The data consists of 12 hours of video-recorded instrumental lessons at the high school and college levels. The lessons have been transcribed in great detail with the aim of representing the actions in three modes: speech,

gestures, and music. Because the original data in the study are in Swedish, a transcribed sequence from a documentary in English (Moore, Connolly, & Anderson, 2001) will be used as an example to illustrate the method and the implications for analysis and interpretations of multimodal video data.

The transcribed scene is from a lesson in composition at tertiary level, typical in its structure within music educational settings in the form that we study. In a one-to-one situation, a female teacher and a female student play the piano and talk about a piece of music composed by the student. The lesson starts with the student presenting her work and ends with the teacher leaving the room with the student in tears. The sequence is chosen for its outlined patterns of interaction, dramaturgically enhanced in a way that highlights patterns found in the study. To facilitate the reading of the complex rationale for the method chosen, the transcription and analytic coding of the example has been attached in [PDF](#) format. The analysis of this sequence will be interpreted further in the article. Due to the amount of detail, it is not possible to include the chart in full format in the article. To view all of the details in the chart on screen, please use the zooming tool in your PDF reader. The chart provides the reader with some insights into how the study is carried out in detail, which, we hope, helps to evaluate the arguments presented in the article. The full transcription and analysis chart of the 5-minute sequence is attached as an appendix to this article.1

Not much is known on a scientific level about interactional dynamics in music instrumental teaching. Nor do we know how the teacher-student interaction shapes the students' opportunities to learn (Hallam, 1998; McPherson, 1993; Rostvall & West, 2003; West & Rostvall, 2003). One of the reasons could well be the traditionally held view that the outcome of music teaching primarily is a consequence of each student's musical aptitude. Many people and even music teachers regard musical talent as a mysterious gift that you either have or do not have. This view has consequences for music education at all levels, which can be illustrated in the following scene in the documentary. In this scene, the teacher says to her student, "There are so many people walking around this earth trying to be composers and they can't bloody write music. Don't be one of them, you have got a lot of talent" (Moore et al., 2001). The interesting point here is that the teacher talks about composing as a talent rather than how the composing skills could be developed.

Another possible reason for the lack of empirical studies on interaction could be embedded in the theoretical and methodological difficulties involved in studying such multifaceted data. Previous research on interaction has been criticized for building on a simple communication model of sender-message-addressee (Shannon & Weaver, 1949) that does not describe the actual functions inherent to complex communication (Eco, 1984; Fairclough, 1995). This model stresses a focus on individual and local descriptive goals that reveal no understanding for higher levels of social institutional formations. The model of initiation-response-feedback proposed by Sinclair and Coulthard (1975) is typical for this critique. Another useful example is the categories of Bellack, Kliebard, Hyman, and Smith (1966), which lock teacher and students into fixed roles and leave little room for a more dynamic view of their interaction. In addition, the speech-act theory proposed by Austin (1962) and Searle (1969) could be criticized for its rigid categorization, which isolates utterances outside their contextual setting.

In the current study, communication is looked on as a series of messages that are divided analytically into message units (J. Green, 1999; J. Green & Dixon, 1994; J. Green, Franquiz, & Dixon, 1997). A message can start in one communicative mode and transcend over to another, from spoken language to gesture and music. Mode is a culturally and socially fashioned resource for representation and communication (Kress, 2003; Kress, Jewitt, Ogborn, & Tsatsarelis, 2001;

Kress & van Leeuwen, 2001). Speech, music, and gesture do not always seem coherent when viewed independently of one another. Each mode has its own semiotics, and the meanings of modes are intertwined and contribute to how instructions can be understood. According to Kress (2003), the "stuff" of our communication needs to be fixed in a mode: Knowledge or information has no outward existence other than in such modal fixing. Each mode has different limitations and potentials. Time-based modes such as speech, dance, gesture, action, and music have logics and potentials for representation that differ from space-based modes, such as image, sculpture, layout, architectural arrangement, and streetscape. Multimodal texts are made up of elements of modes that are based on different logics, and knowledge changes its shape when it is realized in the different modal material (Kress, 2003).

Learning is studied here as a dynamic process of transformative sign making that actively involves both teacher and student. This perspective reveals challenges to the analysis as well as for the designs for learning. An expansion of the access to diverse semiotic resources increases the repertoire of possible thoughts and actions within the field for both teachers and students. This reflects a movement of focus from the traditional division of instruction and construction, toward focusing the educational functions of both instruction and construction in designs for learning.

Video recordings of interaction within a classroom setting generate enormous amounts of multimodal data, and as a result many choices have to be made about how to represent, describe, analyze, and interpret data systematically. Recent technological developments have made it deceptively easy to record human actions and interactions to capture both sound and moving image. A recent review of the literature dealing with the use of video in the social sciences shows how the field has evolved during the past few decades (Rosenstein, 2002). However, it is not easy to handle the dense flow of information in a way that is both systematic and transparent. The analysis of video data is very time consuming. To bring down the amount of information, it is necessary to focus the area of interest by making as many systematic strategic decisions as possible before shooting the video, and certainly before deciding the means of analysis. All systems for analysis are more or less impregnated with assumptions and theories, as they bring different data to the forefront. It is therefore necessary to account for the theoretical background that has led to the methodological decisions resulting in an open-ended software solution for data handling and analysis.

The perspective and design of the ongoing study reflect the researchers' values and views concerning the empirical research field: Instrumental teaching is a complex social phenomenon with a long history. From this perspective, it is problematic to study and discuss the outcome of music teaching merely as a result of teachers' and students' individual aptitude. On the other hand, sociological macro models of explanation or theories about the historical context in which the institutional routines have evolved cannot provide analytical concepts for analysis on a microlevel, in which teachers and students interact.

Our aim in this article is, therefore, to discuss how multiple compatible theories are synergized in the current study for application on different theoretical levels. One outcome of the theoretical and methodological perspective chosen can be bound in the design of a software tool to aid in the transcription and analysis of video data. By applying a multidisciplinary framework, our general understanding of teaching and learning in terms of interaction can thus be enhanced.

Another finding revealed by this specific method of analysis is that students' attention during lessons often is divided. The method enables a systematic analysis that unveils that students have to shift focus frequently between the printed score, complex motor control, auditory and tactile feedback from the instrument, and the teacher's gestures. Teachers' attention is focused primarily

on giving ad hoc instructions concerning the decoding of symbols or on technical aspects. When the teacher's gestures and language give contradictory meanings, the student has to use a larger part of his or her attention to make sense of the situation. This will have negative consequences for the learning process, as students will have less mental capacity available to focus on the issue at hand. The gestures of the teachers are often used to communicate a negative message that is contradictory to the verbal positive instruction or evaluation. By doing so, teachers can state values and feelings that could be regarded controversial if communicated in words. Conflicting messages reduce the students' possibilities to address the problem directly. They still have to make some sense of the situation however, perhaps by blaming themselves for a problem that has never been addressed openly by the teacher.

The design of the method focuses on dynamic aspects of interaction, rather than a static description of the spoken language. The aim is to bring forward a potential for development of alternative interaction strategies, based on research findings rather than unplanned patterns given by tradition.

Research questions

The study is divided into three levels, each with its own set of research questions and theoretical concepts. The differing levels reflect continuous movement, from the close-up description of how teachers and students act and interact, through a systematic analysis of the patterns of interaction, concluding with an interpretation on a macro level of why they are interacting in the way they do. The main object of the study is to transcribe and analyze the processes of teaching and learning systematically to increase our knowledge about how different interaction patterns in the course of instrumental music lessons affect the students' opportunities to learn. With the particular scope of the theoretical and methodological decisions made, such patterns could emerge with coherence and consistency between different modes; from explicit or implicit expectations; oblique or direct information; who controls the definition of the situation, and how; as well as many more. The results of this analysis will be discussed and interpreted within a wider historical and sociological perspective.

We will discuss the design of the method and the design of the software tool developed to facilitate the processes of transcription and analysis of the large amounts of video data. Transcription and analysis of video data is an extremely time-consuming task, and no less so when dealing with 12 hours of videotape. An efficient way of handling the data was needed. The software renders it possible to transcribe and analyze the entire body of video recordings at a microlevel within a reasonable amount of time and has been developed as a direct consequence of the theoretical and methodological decisions in the study. Although there are several dedicated software packages for video analysis available, both commercial and as freeware, the specific conceptual framework led to the decision to put together a new open-ended solution with already available office software.

Theoretical perspectives on designing a study of interaction

One often-discussed weakness in qualitative studies of interaction in teaching and learning settings is the opaque process of analysis and/or categorization of data. There quite often seems to be a gap between the method used and the claims made in the interpretation of data. The analysis that leads to a certain interpretation takes place in a "black box" with few transparent links between the description in the transcript and the interpretation available to the reader. From our theoretical point of view, it is very important to make the process of analysis and categorization

transparent to the reader, so that different studies can be compared and a precise language can be applied.

Applying a multidisciplinary framework

Interaction in education is a complex phenomenon. An analytical matrix from isolated disciplinary fields—such as education, musicology, sociology, or psychology—could not provide all the concepts required for different aspects of the study if the aim is to interpret data and enhance an understanding of interactional dynamics in an educational setting. An interpretation of data applying a single theoretical perspective runs the risk of promising and explaining too much at a level where the concepts cannot provide a logical explanation for the occasionally implicit questions asked. A way of dealing with this is to apply a multidisciplinary theoretical framework on diverse levels of the study, where each level corresponds to a set of research questions. There has, however, to be a logical and theoretical coherence linking the various levels and concepts based on a general theoretical definition of the studied phenomenon identifying which qualities of the phenomenon are to be studied in the chosen method. Theories and results from other fields can be extrapolated and used as an explicit background for the analysis and interpretation. A study of a multimodal interaction in a complex educational setting requires concepts that define a perspective on learning, teaching, music, communication, education, and interaction on a personal, interpersonal, and institutional level. These concepts are interdependent to the extent that it would not be possible to understand the individual without a society, or a society without individuals. From this standpoint it becomes necessary to apply a critical societal perspective that provides us with an understanding of how the institution is confined within the routine actions of the teachers and students. These actions have evolved and gained legitimization throughout the history of the institution.

In this study, all forms of knowledge and action are regarded as a result of multidimensional social processes. Therefore, the focus lies on what forms of knowledge are constructed in the interaction process and how different forms of knowledge are regarded as an expression for the institutions, where knowledge develops and becomes legitimized. This leads to a perspective in which knowledge and learning are understood within several dimensions: for the individual, within a local educational context, on a institutional level where patterns of actions develop and are given certain values, and, finally, for the society at large, in which different institutions compete for resources, legitimization and status.

> A social institution is (amongst other things) an apparatus of verbal interaction, or an "order of discourse" . . . In this perspective, we may regard an institution as a sort of "speech community" with its own particular repertoire of speech events, describable in terms of the sorts of "components" which ethnographic work on speaking has differentiated—settings, participants (their identities and relationships), goals, topics, and so forth. (Fairclough 1995, p. 38)

The three theoretical levels applied in the current study enable the reader to follow the entire research process and the underlying hypothesis and theories concerning the complex phenomena of interaction in a (music) school setting. An institutional perspective serves as an interpretive framework for the study. The lessons are viewed as social encounters and performances (Goffman, 1959/1990) wherein the participants act to create and re-create social orders at different institutional levels by means of communication routines employing speech, music, and gesture (Fairclough, 1995). The actions of the individuals are understood not primarily in terms of individual choices but as routine actions with traditions and a legitimacy stemming from the history of the institution (Berger & Luckmann, 1966/1991; Douglas, 1986).

To address the phenomena on a variety of levels, we apply theories from experimental research on learning and various forms of memorization processes, which provide a frame for understanding how different forms of knowledge develop and internalize into automatized schemas on an individual level (Arbib, 1995; Bartlett, 1932). On the interpersonal level, we use concepts from education, psychology, linguistics, and communication theory to understand how different forms of knowledge are communicated from one person to another, in particular, identifying the factors in the communicative setting that can either support or constrain the learning process on the individual level. On the third level, we interpret the patterns revealed by the analysis and put these into a societal and historical frame of understanding. The notions of institution (Berger & Luckmann, 1966/1991; Douglas, 1986), interaction (Goffman, 1959/1990), discourse (Fairclough, 1995), and power (Fairclough, 1995; Giddens, 1984) are applied to enhance understanding of how different forms of knowledge develop over time and gain legitimacy in society.

Transparency and replication

Based on experiences from a previous study, we decided to transcribe and code the complete data material. This is not often experienced in scientific video microanalysis; mainly because of the sheer amount of information and because the process of transcribing is very time consuming. In previous research on video-recorded interaction, a selection of sequences from the full data range has typically been chosen for detailed analysis. If such choices are not discussed, this could reflect the researchers experiences and assumptions rather than being based on a transparent and systematic method. A method based on such a selection of "interesting sequences" at an early stage in the process runs the risk of involving choices based on a "preverbal analysis" that is not altogether accounted for.

An analysis made more transparently gives the reader the possibility of following the systematic structure of the transcription to comprehend the empirical base for the analysis and the interpretation of the data. This method also opens up the scientific process for criticism and renders it possible to carry out a closer replication of the study.

Descriptive level—Theoretical perspectives on the process of transcribing

On a basic descriptive level, we derive data from microethnographic multimodal transcriptions of speech, gestures, and music (J. Green & Wallat, 1979, 1981). The lessons are video-recorded digitally, and the recordings are stored on computer hard disks. The process of transcribing multimodal communication into written text is a complex task. The linguistic representation of interaction heard and seen in the video recording, and transformed into a written text, forces the researcher to make numerous choices from a theoretical point of view. One important issue to address is what the level of detail should be in the transcription. It is possible to spend endless amounts of time describing all aspects of actions in different modes in great detail even in a very short video sequence. However, it would never be possible to represent the sequence in life-large scale. This, of course, would not even be desirable. The amount of detail and the selection of what should go into the transcript is a choice directed primarily by the aims and research questions in the study. The consistency and coherence of these choices govern the explanatory value and the claims of the study.

The video recordings and the transcript are regarded as "texts" defined in a broad sense as instances of communication (Kress, 2003). The transcript is a text that re-presents an event and is not seen as the event itself. Following this logic, what are re-presented are data constructed by a

researcher for a particular purpose, not just talk or interaction written down objectively. This is one of the reasons behind the decision to describe in detail all of the choices made in the process of transcribing to minimize black boxes of hidden assumptions. Transcripts are incomplete representations, and the manner in which data are represented affects the range of interpretations possible.

Representing message units

The method of transcription focuses on the events during the lessons as a series of communicative messages (J. Green, 1999) in three often overlapping or simultaneously occurring communicative modes: music, speech, and gesture. A transcription technique that represents solely verbal communication is not adequate if one truly claims to analyze the interaction process and the didactic functions of different actions. One could argue that the use of multiple modes is a specific feature of music education, but all forms of communication are mediated through various modes, each one specifying its means of representation and communication (Kress, 2003).

The events are transcribed and analyzed as a series of connected or disconnected communicative messages in three separate communicative modes. Hence, it is necessary to represent all three modes in a coherent way to analyze the events comprehensively. Messages are communicated through message units in more than one single mode, beginning in one and then transcending on to another. Even though we have a great deal of experience from music education, it would not be possible to comprehend what was going on if we did not connect the gestures or the music to the speech. The transcription of the various modes in a single transcript chart made it possible to analyze whether the messages conveyed in the three modes were coherent or incoherent. This technique could reveal inconsistencies and conflicting messages within the communication process. Therefore, the possibility to simultaneously view the transcription of the different modes side by side is of great importance. This makes the graphical layout of the transcription as well as the coding of the modes vital in terms of what can be shown. The representation of the teacher-student dialogue is based on the message units, which are symbolized with a new cell in the transcript chart. The beginnings and endings of the message units are distinguished by "contextualization cues" such as pauses, prosody, gestures, and so on (J. Green, 1999; J. Green & Dixon, 1994; J. Green, Franquiz, et al., 1997).

The musical events and gestures during the lessons are transcribed in the same way, that is, each musical phrase is described in a new cell. The transcript chart is divided into columns: time code, teacher's musical activity, students' musical activity, teacher's talk, students' talk, teacher's gestures, and students' gestures. On the descriptive level, we apply two forms of written representation of the actions on the studied films. In the final version of the data presentation, the multimodal transcription charts are complemented with a chronological narration of the actions to give a picture of the series of the communicative events during the course of the lesson.

Analytical level

On the second level of the study, an analytical framework has been developed based on research concepts from the fields of education, linguistics, cognitive science, and music psychology. We have used the concepts as a matrix that enables us to analyze how different patterns of interaction support or constrain the students' learning possibilities. The interaction is discussed in relation to research on optimal learning conditions in music.

In analogy with critical discourse analysis, we aim to "map three separate forms of analysis onto one another: analysis of (spoken or written) language texts, analysis of discourse practice

(processes of text production, distribution and consumption) and analysis of discursive events as instances of sociocultural practice)" (Fairclough, 1995 p. 2). This is part of an ambition to fulfill the principle "that analysis of texts should not be artificially isolated from analysis of institutional and discoursal practices within which texts are embedded" (p. 9).

The analytical level focuses on discourse and learning processes on an individual and interpersonal level. However, the patterns revealed by the analysis are interpreted not primarily as actions based on each individuals' conscious choices from a set of rational options but as routine actions based on a historically evolved praxis that in many cases has become an opaque structure for the participants (Fleck, 1935/1981; Fairclough, 1995). The opacity renders it difficult for the individual to choose an alternative action if it has not been given legitimacy by the tradition, even though alternative actions are always possible. These patterns have different consequences for individual students' possibilities to learn, which are discussed in relation to research in the fields of cognitive science and music psychology. The example from the documentary film shows how the professor communicates her disapproval through harsh, judgmental comments and intrusive body language. Such an interactional pattern addresses an emotional level, which in the end of the excerpt makes the student cry. The emotional focus forces the student to prioritize her emotional state of mind, which will reduce the potential to grasp cognitively the issue at hand and learn the skills needed.

Analytical concepts

Communication can be analyzed in several ways. This study focuses on the educational functions of each message unit in the transcription. The analytical concepts are selected to focus on how patterns in the interpersonal communication could either support or constrain learning on an individual level. Five educational functions of language, music, and gestures are differentiated with inspiration from sociolinguistics and adapted in the previous study: testing, instructing, accompanying, analytical, and expressive functions. Each transcribed message unit from the teacher and students is coded with the concepts, and the frequencies of the different functions of speech, gesture, and music usage are registered for each lesson, and for teacher and students, respectively.

We have used the cognitive concepts of schema internalization (Arbib, 1995; Bartlett, 1932; Dowling & Harwood, 1986) and focus of attention (Allport, 1980; Bamberger, 1996, 1999; Navon, 1985; Shaffer, 1975; Treisman & Davies, 1973; Treisman & Gelade, 1980) as explanatory models of musical learning at the individual level. The concepts of schemata and focus of attention are differentiated into four categories: cognitive, motor, expressive, and social. Each utterance in the transcript is coded with one of the concepts. The theory of attention has shown that it is difficult to divide the attention on different tasks, such as listening to the teachers talk while trying to get the fingering right as well as playing through a new sequence. Some forms of communication render it difficult for the student to focus on and internalize an adequate motor schema, as most of their attention is geared toward decoding language and the different symbolic systems used by the teacher.

The application of the concepts of schema and focus of attention renders it possible to trace learning sequences (Gordon, 1993) through the observation and coding of actions. By coding each message unit by its educational function and the focus of attention, respectively, we can relate the emerging patterns to previous findings on optimal learning situations.

The analysis shows how the process of schema internalization is facilitated or restrained when the focus of attention changes during the interaction between teacher and student. One outcome of

the analysis is that the understanding of the situations requires that the student and teacher agree on what to focus on to develop cognitive, motor, expressive, and social schemas. If the teacher prematurely forces the students to focus their attention on a level that is not yet accessible for the student, the teaching will delay the schema internalization rather than support it.

We also analyze whether the use of educational genres in the different modes are coherent during critical sequences. The analysis shows numerous examples of situations in which inconsistency in the use of various modes created problems for students in grasping the task at hand. One example of incoherent messages was the very frequent teacher utterance, "Very good, take it once more." The positive verbal comment was regularly accompanied with an incoherent message that was communicated with gestures. The teacher showed his disapproval with the student's manner of playing by his way of moving in the classroom. While students are emotionally occupied trying to interpret the incongruous emotional message, they will have less mental capacity to focus on the issue at hand. Their attention has to be divided between different tasks, which can delay the schema internalization process.

The situations described above are interesting to analyze in detail, as they demonstrate critical moments in teaching and learning. Analysis of these situations can help us to understand how the multimodal communication in the interaction affects students' possibilities of learning. Another critical factor in the interaction and learning is the teachers' mode of using language to accompany their instruction in other modes. In the previous study, we concluded that many misunderstandings occurred when teachers used language inconsistently. The guitar teachers in the previous study frequently mixed four to five symbolic systems for fingering and counting the rhythm while they were trying to help their student to play through a new sequence.

Metalevel

On the third level, the results from the description and analysis on individual and interpersonal levels are discussed in terms of their institutional and societal origins and how these patterns affect the distribution of musical knowledge in society in terms of power and hierarchy. The metalevel is the third stage in mapping the interaction and their consequences for students' possibilities to learn (Fairclough, 1995).

From a perspective of institutional theories (Berger & Luckmann, 1966/1991; Douglas, 1986), the actions of the individuals are understood not primarily as results of individual choices but as routine actions with traditions and legitimization inherited as a part of the history of the institution. The patterns of interaction that have evolved during the institutional history is, to a large extent, defining what is possible to achieve in today's classrooms, and this, consequently, reduces the individual teacher's ability to choose his or her actions freely. A perspective that exceeds the individual level—like a historical background based on earlier research in musicology—is, hence, essential as a framework to show the historical development of the institutions that have shaped and influenced the studied activities.

On the metalevel, the results from the analysis on the personal and interpersonal levels are interpreted so that we can understand the patterns of interaction. Coding each message unit with the analytical concepts gives a picture of patterns of interaction both typical and atypical for the institutions. In the previous study, we found, for example, that teachers very rarely focused on the expressive features of the music played, even though this could have been expected, as the official guidelines for the municipal music schools emphasize expressive features.

The emerging patterns are critically viewed to increase our understanding of the distribution of knowledge and power. The patterns are compared to previous findings on optimal learning conditions. It is also of importance to compare the ideological discourse on education within the institution and in the surrounding society against what is observed to take place in the monitored lessons. One interesting finding has been an obvious gap between the discourse at an official level and the observed actions in the classroom (Rostvall & West, 2003).

Historical background

In Western music tradition, music is regarded as materialized in the printed score and is supposed to exist as an object even when it is not played. Musicians, composers, and conductors have different roles in a hierarchical order. The roles of making and performing music are separated. Composing and directing are regarded as highly differentiated skills based on special talents that only a few possess. The musician has to act on composers' and conductors' ideas rather than expressing him- or herself. In the same hierarchical manner of regarding music, highly skilled musicians and/or instrumental teachers are viewed as "masters" who expect and receive respect and obedience from their students. The music teacher is regarded as gatekeeper of the tradition, and if the student shows him- or herself worthy of it, the master can share the tradition with him or her, but only a small piece at a time. This is the manner in which musical knowledge is presented within the method books as very short sequences, often only one note at a time.

In instrumental teaching, music is typically treated within the tradition of objectification. This notion of music as artifacts in "the imaginary museum of musical works" (Goehr, 1992) affects the way in which popular music, music making, and music as a means of action are treated (Elliott, 1995; L. Green, 1997). According to the ideals of this tradition, the musical object—the piece—should be presented to the student, who, after some practice, should be able to play the music in the manner approved by the institution. If students do not achieve this, they are almost immediately regarded as untalented or unmotivated. If we return to the example from the documentary film (Moore et al., 2001), the analysis exposes how the notions of musical talent and the hierarchical order are carried out in the interaction between the student and her professor.

By using harsh and judgmental language that refers to the student as an individual, rather than explaining how she can develop her skills, the teacher in the example controls the situation and remains the master. The means of using and evoking strong emotions in the communication renders it very difficult for the student to confront the teacher on a cognitive level, for example by questioning why the teacher does not tell her how she should study and what she is expected to accomplish. The teacher has given a contradictory task that, consequently, cannot be solved. The student was asked to compose a piece in a historical style. When she then presents her piece in the style of Palestrina, the teacher tells her that she has failed in her task, because the piece sounds too much like Palestrina's work. The professor never mentions how she comes to that conclusion or how the student should think when she works with such a task to fulfill the expectation. By being indistinct in her manner of communication, the teacher can convey that the student should solve problems intuitively. The result of the student's work will then be judged on a level without words, and is impossible to challenge. This communication pattern upholds the notion of talent and obedience in the institution of music teaching.

The pedagogical consequences of such a traditional view are that the communication between teacher and student does not need to focus on how the student should develop desired competences. Instead, the main focus is on the teacher's often nonverbal positive or negative judgment of the student's process of trial and error. The outcome of the tuition is then a result of how well the student can perceive and process nonverbal information without actually being

taught explicitly how to achieve specific skills. From this rather conventional point of view, the outcome of music education is dependent on two individuals' static talents, without considering communicative conditions. To explain human actions in situations like this, an understanding of the historically shaped institutional patterns of interaction and distribution of power needs to be added.

A common problem in instrumental music education in Sweden is that many students drop out during their first year. This is typically explained as a result of individual reasons, such as a lack of musical aptitude or motivation. Other explanatory models can, however, reveal further causes for common problems in education. The asymmetric interaction could be reason enough for students ending their tuition. Their other options are to accept or to challenge the teacher's preferential right to define the situation (Goffman, 1959/1990) and thereby run the risk of getting involved in an open conflict with the teacher. The notion of power is essential in the understanding of interaction (Giddens, 1984; Goffman, 1959/1990). Our emphasis is placed on the distribution of power and on who is to control the definition of the situation.

The communicative patterns during the lessons have great influence on the students' learning possibilities. Incongruent messages, imprecise use of language, and a method of instruction that forces the students to divide their attention between too many different tasks will reduce the student's possibilities of learning rather than support the learning process. If the tuition does not give the students all of the experience and all of the cognitive skills they need, they become dependent on other sources of information to be able to fulfill the expectations of the teacher. The example from the documentary illustrates this, since the teacher never explains how the student should work to develop her skills. The professor only tells her to persist in her efforts by starting all over again.

Instead of acknowledging a problem directly, teachers very often interrupt the students and tell them to play once more without giving sufficient verbal information. To be able to succeed with the tasks given during tuition, the students have to have other opportunities to learn outside of the classroom. Students without relatives that play an instrument often have not had enough prior access to musical experience to complement the instrumental tuition. This could have an effect on their ability to comprehend the fragmentary use of language and music observed, as well as the musical notation in the method books. Students without adequate information or supplementary experience run a high risk of failing and could come to be regarded as untalented or less motivated. The perspective and method used in this study reveal and address the consequence of failure on the interpersonal and institutional levels, whereby teachers, schools, and authorities can respond rather than seeing the individuals as being personally responsible.

The institutional perspective can also be discussed from an ethical point of view. Video analysis of interaction reveals events that are not obvious to the participants, as they are fully occupied coping with the situation. Such revealing data could cause a lot of anxiety on the individual participant if the explanatory models used in the analysis focus primarily on the individual performances of teachers and students.

Methodology and methods

An ambition in the project is to make the entire research process transparent to facilitate critical reading, as well as a reproduction of the study. This also reflects an ethical ambition to aid in following the transformation of the actions of the informants through the processes of transcription and analysis, on to the presentation of results and discussion of implication for practitioners.

The question of how different patterns of interaction affect the student's opportunities to learn is discussed on the analytical as well as on an interpretative metalevel. The Analyzing and Reporting Transcription Tool—ARTT—makes this feasible even with very large amounts of video data.

In this project, we were challenged with finding a way to handle the large amounts of video data systematically and transparently using a combination of commonly used office software. The software tool developed rendered it possible to view the digital video synchronized to a spreadsheet containing connected fields for transcription and coding interaction in different modes, all on a single computer screen. The spreadsheet is programmed to aid the recognition of patterns in the interaction, as well as more quantitative modes of output.

Instrumental lessons are videotaped using a digital camcorder. The recorded lessons are then transcribed and analyzed in full at a detailed microlevel. Participation in the study is voluntary, and the informants are well informed about how the video material will be used. Written consent is required of the participants or their parents, if they are under 18 years of age.

The headmasters of each participating school arrange contacts with teachers willing to participate. After thorough explanation, the teachers discuss further with his or her students, and distribute information and gain their agreement and that of their parents. The families can access further information from the project Web site (http://www.didaktikdesign.nu/musik) before signing the agreement to participate. After scheduling the fieldwork, a research assistant visits the school to capture the lesson in its natural setting. A small digital (mini-DV) camcorder (Panasonic NV-MX 350) with a wide-angle lens and a built-in microphone of good quality is placed on a tripod so that it will capture both the teacher and the student(s). Additional information about the room is noted by the assistant and registered by a full-circle camera sweep. The assistant leaves the informants alone in the room to reduce the effect on the tuition to a minimum. Immediately after the lesson is over, the participants are invited to watch the video on a laptop computer (Macintosh PowerBook G4) running video software (Apple iMovie). The computer is hooked up to the camcorder through a high-speed connection (FireWire, also known as iLink, DV-out, or IEEE 1394 Standard), so that the video can be run directly from the camcorder. After viewing the lesson, the informants have the opportunity to opt out of the study, in which case the tape is immediately erased. A small questionnaire captures background data from the participants. An ethical decision to keep the informants anonymous, to comply with the standard of ethics issued by the Swedish Research Council, led to our decision to shoot three times the amount of video that was needed and to make a randomized sample of the lessons to be analyzed. This is done in an effort to protect the identities of the informants further.

The large amounts of data generated, together with a detailed systematic representation, description, analysis, and interpretation, have created the need for efficient data handling. The analyzing and reporting transcription tool developed consists of two commonly used software packages, connected through a third software package that makes it possible to program simple strings of code to control the system software to enable the different software programs to interact.

The digital video runs in Apple QuickTime Player, and the transcript chart runs in a spreadsheet on Microsoft Excel. Figure 1 shows a screen shot of the system. To protect the identities of the informants, the content is not from the study—the picture is a commercial image published with permission to illustrate the software (© Royalty-Free/Corbis).

The spreadsheet contains connected fields for transcribing and coding separate modes of interaction. It is programmed to facilitate the recognition of patterns in the interaction by assigning a color to each of the analytical concepts used for coding each utterance. The possibility of viewing different modes next to each other gives an overview that graphically reveals patterns in the interaction. Quantitative calculations are fairly simple to accomplish, as this is what spreadsheets are designed to do. It can easily be programmed to make automatic statistical calculations, for example of durations and frequencies, as well as referential and correlation analysis, of several aspects of interaction. Output from such calculations and analysis can be displayed in several types of graphs. The Excel spreadsheet is programmed so that it can calculate the time code format with hours, minutes, seconds, and frames per second. This is done with an extension file. The whole transcript sheet with the coding fields contains so many fields with so much information that it would be hard to get an overview or even to fit on a standard size computer screen. Because the need for viewing all fields in the complex transcript and coding chart at the same time is limited, the chart is programmed to display combinations of fields needed for different tasks while hiding the other fields. Keyboard commandos programmed as macros within the Excel software display these different views (Figure 2).
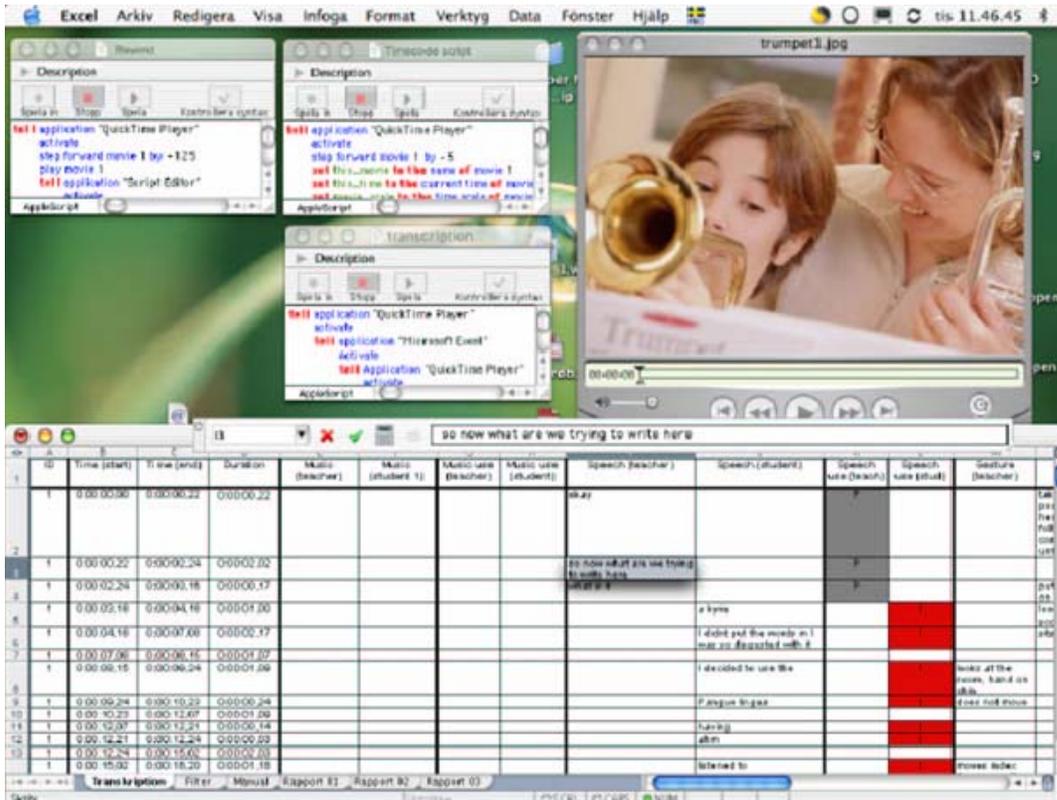
Figure One

Music (teacher)	Music (student)	Music use (teacher)	Music use (student)	Speech (teacher)	Speech (student)	Speech use (teach)	Speech use (stud)	Gesture (teacher)	Gesture (student)	Gesture use (teach)	Gesture use (stud)	Theme (teacher)	Theme (student)
				okay		P		takes out a portfolio from her bag, eyes following the content as she unfolds it			T	K	K/S
				so now what are we trying to write here		P					T	K	K/S
				what is it		P		puts the score on the table			T	K	K/S
					a kyrie		T	looks at the score			P	K	K/S
					I didnt put the words in I was so disgusted with it		T	sits very still			P	K	K/S
					I decided to use the		T	looks at the score, hand on chin		P		K	K
					Pangue lingua			does not move		P		K	K
					having		T			P		K	K
					ahm		T			P		K	K
					listened to		T	moves index finger over the mouth		P		K	K
					Josquins							K	K
				(sigh)		T		puts down the hand			T	K	
				see		T		leans forward			T	K	
				what		T		rises from chair			T	K	
					pangue lingua mass		T	passes in front of the student			T	K	
				yeah		T		turns the back to the student and continues towards the piano			T	K	
				i hate octaves		T					T	K/S	K/S
				karen why do		T			rises from chair	T	P	K/S	K/S
				why do you write		T		"flips" the score on to the music stand on the piano	follows behind the teacher	T	P	K/S	K/S
				which paper wouldnt use octave		T			continues up to the teacher at the piano	T	P	K/S	K/S
				only		T		sits down at the piano	moves her left hand through her hair, looking at the score	T	P	K	K/S
				occasionally		T		looks at the score		P	P	K	K/S
				after justifying		T		lifts her right hand over the keyboard	stands behind the teacher's right shoulder, right hand on the...	P	P	K	K/S

Figure Two

Both QuickTime and Excel are controlled through AppleScript, a simple programming language used to write script files for the Macintosh operative system, which minimizes keyboard actions on the computer and the applications that run on it. Scripts are used to facilitate transcription and coding in several ways. One script lets the QuickTime Player be controlled with functions similar to a dictating machine, for example starting the video at an earlier point to where it was last stopped. A more complex script changes the format of the video time code, copies that code to the computer's clipboard, pastes it into the active cell in the Excel spreadsheet, then starts an Excel macro that activates the following cell, and finally backs the video to play the previous three seconds before continuing to play, all at a single click on the mouse. This script also compensates for the delay in response time of the operator.

The video is imported to the computer from the camcorder with the Apple iMovie software, an easy-to-use video-editing program that automatically detects if the camera is connected to the computer. Video data require large amounts of memory and storage space in the computer, and to be able to work with all of the video material in a single data file, Apple QuickTime Pro is used to save the video in the compressed MPEG-4 format. This renders it possible to have the entire video material connected to a single spreadsheet, accessible for both qualitative and quantitative analysis. These large files can be stored together on an external hard disc or on a single DVD.

All of the software is commonly used and requires a minimum of additional programming. Microsoft Excel is a part of the standard Microsoft Office package; QuickTime Player is a free standard program for running video on the computer. These are available for a variety of operating systems, including Microsoft Windows and Apple Macintosh. AppleScript and iMovie are preinstalled in the Macintosh system. This also keeps the need for support and training to a minimum. An important reason for developing a tool for analysis using widespread office software, rather than using a dedicated video analysis software package, is to make research-

related methods transparent and accessible for teacher-training programs and professional development programs in schools, as well as for practitioners and the wider public.

Analytic Concept	Letter	Color
Instructing	I	Red
Accompanying	B	Deep blue
Testing	P	Dark gray
Analytical	A	Dark green
Expressive	E	Dark yellow
Cognitive	K	Light blue
Motor	M	Orange
Social	S	Light gray
Cognitive and expressive	L	Light green
Cognitive and motor	H	Purple
Cognitive and social	O	Dark pink
Expressive and motor	W	Light yellow
Expressive and social	Q	Light pink
Motor and social	T	Light red

Transcription

The first level consists of 12 hours of videotape, the transcription charts of the multimodal communication, and a chronological narration of the events and actions during the lessons (Figure 3).

rt)	Music (teacher)	Music (student)	Speech (teacher)	Speech (student)	Gesture (teacher)	Gesture (student)	N
0			okay			takes out a portfolio from her bag, eyes following the content as she unfolds it	only student v
2			so now what are we trying to write here				
4			what is it			puts the score on the table	
6				a kyrie		looks at the score	
6				I didnt put the words in I was so disgusted with it		sits very still	
8							clip
5				I decided to use the	looks at the score, hand on chin		only teacher's
4				Pangue lingua	does not move		
3							
7				having			
1				ahm			
4							
2				listened to	moves index finger over the mouth		
0				Josquins			
7			(sigh)		puts down the hand		loud inhalatio
2			see		leans forward		
1			what		rises from chair		
1				pangue lingua mass	passes in front of the student		camera follow
7			yeah		turns the back to the student and continues towards the piano		both visible
5			I hate octaves				
8			karen why do			rises from chair	
9			why do you write		"flips" the score on to the music stand on the piano	follows behind the teacher	
4			which paper wouldnt use octave			continues up to the teacher at the piano	teacher hidde student
1			only		sits down at the piano	moves her left hand through her hair, looking at the score	
0			occasionally		looks at the score		both visible
4			after justifying		lifts her right hand over the keyboard	stands behind the teacher's right shoulder, right hand on the hip	

Figure Three

The vertical columns in this view of the spreadsheet are used to register speech, music, and gesture. In the horizontal rows, these actions are divided into communicative units, or utterances separated by changes in prosody, gestures, or music (J. Green & Wallat, 1979, 1981). Time is registered for each utterance.

A narrative description of the tuition also contains extracts from the transcriptions. The combination of different forms of representation enables the reader to follow parallel actions in different modes: teacher and student speech as well as actual music performance, in addition to other forms of actions such as gestures and eye contact.

Analysis

On the second level of study, we analyzed the transcriptions of the multimodal communication using theoretical concepts to differentiate and systematize patterns of interaction in the teaching and learning process. The analytical concepts in the study are developed from educational genres of speech and music usage (Rostvall & West, 2003). This provides a multimodal analysis of the use of speech, music, and gesture, as well as method books, according to a perspective based on traditions and needs in the specific music-educational setting (Figure 4).

Time (start)	Speech (teacher)	Speech (student)	Speech use (teach)	Speech use (stud)	Notes
0:00:00,00	okay		P		only student visible
0:00:00,22	so now what are we trying to write here		P		
0:00:02,24	what is it		P		
0:00:03,16		a kyrie		I	
0:00:04,16		I didnt put the words in I was so disgusted with it		I	
0:00:07,08					clip
0:00:08,15		I decided to use the		I	only teacher's face visible
0:00:09,24		Pangue lingua		I	
0:00:10,23					
0:00:12,07		having		I	
0:00:12,21		ahm		I	
0:00:12,24					
0:00:15,02		listened to		I	
0:00:16,20		Josquins		I	
0:00:17,07	(sigh)		I		loud inhalation
0:00:17,22	see		I		
0:00:18,21	what		I		
0:00:20,01		pangue lingua mass		I	camera follows the teacher
0:00:21,07	yeah		I		both visible
0:00:21,15	i hate octaves		I		
0:00:22,18	karen why do		I		
0:00:22,19	why do you write		I		
0:00:23,24	which paper wouldnt use octave		I		teacher hidden behind student
0:00:25,21	only		I		
0:00:26,10	occasionally		I		both visible
0:00:27,04	after justifying		I		

Figure Four

Separate educational functions of speech, music, and gesture during the lessons are differentiated. On a level of detail where units typically are shorter than a second, each utterance is coded with one function only. Of course, all interactions contain several levels of different meanings; our aim, however, was to achieve such a level of detail that we could discern where the results reveal

themselves. Those small message units must be coded only one way; otherwise, they would have had to be divided into two separate units. The coding had hardly been possible without deciding what to look for. This coding specifically deals with the educational functions of utterances, that is how they affect opportunities for learning. In this view of the transcription chart the columns where the speech utterances and the coding according to these functions appear concern teacher and student respectively. The coding cells have an automatic function that assigns a color to each of the functions. This facilitates rapid recognition of patterns in the usage of music, speech, and gesture.

The view in Figure 5 displays the coding of the music, speech, and gesture usage for the teacher as well as the student. In this particular sequence, the teacher uses a mode of speech and gesture coded as testing, whereas the student mainly uses an instructive mode of gesture and speech. The columns for music usage are empty, because the sequence contains no musical utterances (Figure 6).

Figure Five

Figure Six

There are also columns in the transcript for describing where the teacher and the student respectively focus their attention (Shaffer, 1975; Treisman & Davies, 1973). In combination with cognitive concepts of experiencing and learning music by developing internal schemata (Arbib, 1995; Bartlett, 1932), we compare the focusing of attention on different concrete targets and different forms of knowledge during the music lessons, with varying ways of using language, music, and gesture in interaction. The color coding of the cells in the transcript also facilitates these comparisons. In this particular view the fields containing the coding of music use, speech use, gesture use, and focus of attention are visible for teacher and student, respectively (Figure 7).

0:00:09,24	0:00:10,23	0:00:00,24
0:00:10,23	0:00:12,07	0:00:01,09
0:00:12,07	0:00:12,21	0:00:00,14
0:00:12,21	0:00:12,24	0:00:00,03
0:00:12,24	0:00:15,02	0:00:02,03

Figure Seven

Further software development

In parallel to the study, efforts have been made to make the process of transcription and coding still more efficient. The largest obstacles with the system used were that problems sometimes occurred when transferring data between computers and that the system did not work fully on platforms other than the Apple Macintosh. In addition, an interest from other researchers in using the system showed that it took more-than-average knowledge to modify the Excel sheet for use with other sets of data. The process of developing a cross-platform solution with a less complex user interface has now resulted in a single java-based application called the Video Analyzer (West et al., 2005). This application has been developed in close collaboration with computer programmers and interaction designers, and is available free over the Internet at http://www.sonartdesign.se. The software is open source, to invite others to take part in the development. Compiled code, source code, and comprehensive documentation are, therefore, also available for downloading.

The Video Analyzer works with Windows and Macintosh platforms, and integrates transcription, coding, and video synchronization in an easy-to-use interface. The design of the application separates the logic of the system, the actual data, and the presentation. The transcripts are saved in xml format readable by several other applications, including database and statistical analysis software for further work or presentation. Data are transferable over a network, which opens up a potential for collaborative work on shared sets of data, interchange of findings, and triangulation of analysis, to name but a few possibilities.

Toward a multidisciplinary perspective on studies of interaction

In many instances where a qualitative perspective is used in studies of interaction and education, there seems to be a gap between a descriptive level and the claims being made about implications for educational practice. A multidisciplinary perspective could fill the gap between descriptions of educational interaction and the claims being made when analyzing and interpreting data from educational settings. Such studies gain by employing different, yet compatible, theoretical levels to describe, analyze, and interpret the complex and historically evolved interaction patterns as

well as the relationship between teacher-student interactions with regard to opportunities for learning.

This should not be seen as a criticism of current theories and methods in educational research but, rather as a critique of claims that scarcely can be drawn when employing any single method. Different theoretical perspectives each have their own field of explanatory value, and a critical eye should look for instances where no argumentation connects the descriptive level of a study with an analytically based interpretation of the data. Such black boxes could, of course, contain a logical link between the different levels; the point is that this should be as open to the reader and explicit as possible.

To go beyond the descriptive level of the complex events in an educational situation, we need to ask questions about what is happening in the situation, how such events are formed, and why these actions occur rather than others. A perspective on these questions at different yet connected levels of study could usefully apply diverse ways of studying them. As long as the different theories and methods used are compatible, they could give consistency between the different levels of the study that brings about deeper and broader understanding. This is not to be confused with a random eclectic perspective, whereby different aspects that fit with the claims are collected more or less by convenience. There is a logical and common ground between the theories of schema, institution, and interaction used in the described study, as they are coherent models that provide an understanding of networks of experiences and actions, at different individual and societal levels.

The consistency and compatibility of the different theories and perspectives that are used on different theoretical levels in the study described is designed to give a deeper understanding of teaching situations on both a personal and interpersonal, as well as on a societal level. This brings with it a potential for change and development that could lead to a structured redesign of tuition based on what we know about interaction and learning.

Notes

1. Please refer to the Methodology and Methods section of this article for an explanation of the transcription chart.

References

Allport, D. A. (1980). Attention and performance. In G. Claxton (Ed.), *Cognitive psychology: New directions* (pp. 112-153). London: Routledge Kegan Paul.

Arbib, M. A. (1995). Schema theory. In M. A. Arbib (Ed.), *Brain theory and neural networks* (pp. 830-834). Cambridge, MA: MIT Press.

Austin, J. L. (1962). *How to do things with words: The William James lectures delivered at Harvard University in 1955*. Oxford, UK: Clarendon.

Bamberger, J. (1996). Turning music theory on its ear. *International Journal of Computers for Mathematical Learning*, *1*(1), 33-55.

Bamberger, J. (1999). Learning from the children we teach. *Bulletin of the Council for Research in Music Education*, *142*, 48-74.

Bartlett, F. C. (1932). *Remembering*. Cambridge, UK: Cambridge University Press.

Bellack, A. A., Kliebard, H. M., Hyman, R. T., & Smith, F. L. (1966). *The language of the classroom*. New York: Teachers College Press.

Berger, P., & Luckmann, T. (1991). *The social construction of reality: A treatise in the sociology of knowledge*. London: Penguin. (Original work published 1966)

Douglas, M. (1986). *How institutions think*. New York: Syracuse University Press.

Dowling, W. J., & Harwood, D. L. (1986). *Music cognition*. Orlando, FL: Academic Press.

Eco, U. (1984). *The Role of the reader*. Bloomington: Indiana University Press.

Elliott, D. J. (1995). *Music matters: A new philosophy of music education*. New York: Oxford University Press.

Fairclough, N. (1995). *Critical discourse analysis: The critical study of language*. London: Longman.

Fleck, L. (1981). *Genesis and development of a scientific fact* (T. J. Trenn & R. K. Merton, Eds.; F. Bradley & T. J. Trenn, Trans.). Chicago: University of Chicago Press. (Original work published in German in 1935)

Giddens, A. (1984). *The constitution of society*. Cambridge, UK: Polity.

Goehr, L. (1992). *The imaginary museum of musical works: An essay in the philosophy of music*. Oxford, UK: Oxford University Press.

Goffman, E. (1990). *The presentation of self in everyday life*. London: Penguin. (Original work published 1959)

Gordon, E. (1993). Learning sequences in music: Skill, content, and patterns. Chicago: GIA. (Original work published 1980)

Green, J. L. (1999, September). *Transcribing as a conceptual process: Exploring ways of representing classroom activity*. Paper presented at a workshop, Uppsala University, Institution of Pedagogy, Sweden.

Green J. L., & Dixon C. (1994). The social construction of classroom life. In A. C. Purvis (Ed.), *Encyclopedia of English studies and the language arts* (pp. 1075-1078). New York: Scholastic.

Green, J. L, Franquiz, M., & Dixon, C. (1997). The myth of the objective transcript: Transcribing as a situated act. *TESOL Quarterly*, *31*(1), 172-176.

Green, J. L., & Wallat, C. (1979). What is an instructional context?: An exploratory analysis of conversational shifts across time. In O. Garnica & M. King (Eds.), *Language, children, and society* (pp. 159-174). New York: Pergamon.

Green, J. L., & Wallat, C. (Eds.). (1981). *Ethnography and language in educational settings*. Norwood, NJ: Ablex.

Green, L. (1997). *Music, gender, education*. Cambridge, UK: Cambridge University Press.

Hallam, S. (1998). *Instrumental teaching: A practical guide to better teaching and learning*. Oxford, UK: Heinemann.

Kress, G. (2003). *Literacy in the new media age*. London: Routledge.

Kress, G., Jewitt, C., Ogborn, J., & Tsatsarelis, C. (2001). *Multimodal teaching and learning: The rhetorics of the science classroom*. London: Continuum.

Kress, G., & van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. London: Arnold.

McPherson, G. E. (1993). *Factors and abilities influencing the development of visual, aural and creative performance skills in music and their educational implications*. Unpublished doctoral dissertation, University of Sydney, Australia.

Moore, S. (Executive Producer), Connolly, B. (Director), & Anderson, R. (Director). (2001). *Facing the music* [Film]. Lindfield, NSW, Australia: Film Australia.

Navon, D. (1985). Attention division or attention sharing. I. M. I. Posener & O. S. M. Marin (Eds.), *Attention and performance* (Vol. 11). Hillsdale, NJ: Lawrence Erlbaum.

Rosenstein, B. (2002). Video use in social science research and program evaluation. *International Journal of Qualitative Methods*, *1*(3), Article 2. Retrieved September 21, 2004, from http://www.ualberta.ca/~ijqm/

Rostvall, A.-L., & West, T. (2003). Analysis of interaction and learning in instrumental teaching. *Music Education Research*, *5*(3), 213-226.

Searle J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge, UK: Cambridge University Press.

Shaffer, L. H. (1975). Multiple attention in continuous verbal tasks. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance* (Vol. 5, pp. 157-167). London: Academic Press.

Shannon, C., & Weaver, W. (1949). *The mathematical theory of communication*. Urbana: University of Illinois Press.

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

Treisman, A., & Davies, A. (1973). Divided attention to ear and eye. In S. Kornblum (Ed.), *Attention and Performance* (Vol. 4, pp. 101-117). London: Academic Press.

Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, *12*, 97-136.

West, T., & Rostvall, A.-L. (2003). A study of interaction and learning in instrumental teaching. *International Journal of Research in Music Education*, *40*, 14-26.

West, T., Rostvall, A.-L., Johansson, B., Hermodsson, E., Galaz, V., & Mohammadi, A. (2005). *Video analyzer* [Software]. Retrieved June 6, 2005, from http://sonartdesign.se/applications/video_analyzer_en.html