It's Not Only What You Say, But How You Say It: Investigating the Potential of Prosodic Analysis as a Method to Study Teacher's Talk

Raija Hämäläinen[a], Bram De Wever[b], Teija Waaramaa[c], Anne-Maria Laukkanen[c], & Joni Lämsä[a]

[A] University of Jyväskylä, Finland

[B] Ghent University, Belgium

[C] University of Tampere, Finland

Abstract

In this study, we introduce new insights into prosodic analyses as an emerging method to study what happens in classrooms interactions. We claim that the prosodic aspects (features of speech such as intonation, volume and pace) of talk are important, but under-represented in the learning sciences. These prosodic aspects may be used to complement, intensify or even reverse the linguistic content of speech. Thus far, most research on classrooms has focused on the content (what is said) rather than on understanding the meaning of the prosodic features (how it is said) of talk. In this study, we introduce prosodic analyses as a method to study classroom discussions. Our exploratory experiment focuses on the prosodic perspective of teacher's talk to shed light on classrooms interactions. We present a case in which we align prosodic features with the content of teacher's talk during a nine-week physics course. This article shows that prosodic analyses may have added value for research on learning and professional development. Namely, we illustrate that acting in an authentic classroom setting might trigger specific prosodic aspects in teacher's talk. We further found indications that the teacher applied different voice prosody regarding certain patterns of classroom talk. For the future, we suggest that a combination of content and prosodic analysis is a promising tool for gaining new insights into classroom interactions.

Keywords: teacher's talk; prosodic analyse

Raija Hämäläinen, PhD, Professor, raija.h.hamalainen@jyu.fi   Department of Education, P.O.Box 35, FIN-40014, University of Jyväskylä, Finland DOI: **https://doi.org/10.14786/flr.v6i3.372**

1. Introduction

Multiple methods and techniques are required to understand what happens in classrooms, and while many researchers have investigated the content of talk – for example, with content analysis (De Wever et al., 2006; Hämäläinen & De Wever, 2013) and discourse or conversation analysis (Mercer & Dawes, 2014; Warwick, Vrikki, Vermunt, Mercer, & van Halem, 2016) – few researchers have attempted to understand the prosodic features of talk (Gweon et al., 2013). The present study methodologically bridges the gap between two research domains to advance research on classroom discussions via the analysis of prosodic features. This analysis focuses on elements such as intonation and pitch on teacher's talk in an authentic classroom context from a sociocultural perspective.

1.1 Teachers' Talk in the Classroom: Educational Dialogues and Teacher Monologues

According to Mercer, Dawes, and Staarman (2009), authentic classroom situations typically involve non-dialectic teacher monologues and educational dialogues. In non-dialectic situations, typically only the teacher talks (*teacher monologue*). In classroom contexts, in addition to educational dialogues, non-dialectic situations may be necessary and an intriguing way to stimulate learning. On the other hand, nowadays, the role of the teacher is changing from only providing knowledge to also supporting students' knowledge construction activities. During educational dialogue – also referred to as *productive classroom talk*, which has an analogous meaning (Muhonen et al., 2017) – both students and teachers talk. According to Mercer (1995), classrooms set up 'sceneries' of educational dialogues where, ideally, teachers and their students will collaboratively discuss the topic about which they are learning. In these educational dialogues, teachers engage their students in discussions that include a series of questions and answers.

For sociocultural research, the creation of meaning is inherently an intrapersonal process, and ways of thinking are embedded in particular ways of using language (Littleton & Mercer, 2013). Dialogue may thus be said to be more than 'just talk' (O'Connor & Michaels, 2007). Dialogue is talk that is productive and, therefore, should be the central interest of analysis (grounded on the work of Vygotsky, 1987). According to Muhonen et al. (2017), previous studies on learning have indicated that the quality of teacher-student dialogue is associated with the growth of students' understanding (Alexander, 2001; Lemke, 1990; Mortimer & Scott, 2003). As a direct result, the field particularly needs to understand teacher talk during educational dialogues that creates opportunities to promote learning (see also Mortimer & Scott, 2003). Educational dialogues are influenced first by differential power relations between teachers and students (Lemke, 1990) and second by differential knowledge relations, for example by being either the 'primary' knower (typically a teacher) or a 'secondary' knower (typically a student) regarding the topic under discussion (Berry, 1981). The teacher also plays a special role in guaranteeing that students benefit from classroom activities (Nassaji & Wells, 2000). Some scholars have argued that in (science) classrooms, educational dialogues are likely to follow triadic dialogue patterns (Lemke, 1990), especially during whole-class discussions (Lemke, 1990; Mehan, 1978; Mortimer & Scott, 2003; Salloum & BouJaoude, 2017). For example, according to Nassaji and Wells (2000), educational dialogues typically proceed along an initiation-response-feedback (I-R-F) pattern, which includes three phases. First, the teacher initiates a question (usually with a known answer); second, one or more students respond to that question; third, the teacher evaluates the answers, provides feedback and may or may not ask for follow-up questions or activities. Wells (1993) further highlights that the students' responses are a crucial element of I-R-F, since without such responses, there is no exchange (dialogue). Sinclair and Coulthard (1975) have also called the third stage (feedback) 'follow-up' and further define three types of follow-up acts: (1) accept or reject, (2) evaluate and (3) comment, which includes exemplifying, expanding, and justifying.

1.2 Acoustic Speech Research

From the research area of acoustic speech and voice research, it is known that prosodic features affect speech (later referred to as *talk*) perception and discussion. By *prosodic features*, we mean vocal characteristics like pitch variation and stress pattern, pausing, tempo, mean pitch and loudness, and vocal quality. *Prosody* refers to the intentional or unintentional use of these characteristics to convey the meaning of an utterance (see

the links in the appendix for more information on prosody). For example, prosody may signal one's psychophysiological activity level and emotions. Changes in pitch and/or loudness and tempo may reflect changes in activity or arousal level (Vilkman & Manninen, 1986; Laukkanen et al., 1997; Waaramaa et al., 2010; Waaramaa et al., 2014). Activity or arousal level can be low, moderate, or high. Typically, a positive emotion of joy and a negative emotion of anger have a high arousal level, while tenderness and sadness have a low arousal level. A high arousal level is typically expressed by high pitch and loudness and a firmer (more pressed, tense) voice quality (Laukkanen et al., 1997; Waaramaa et al., 2010). In a low arousal level, the mean pitch and loudness are lower, and the voice quality is less firm (softer, laxer). The *valence* of the emotion (i.e. whether it is positive, negative, or neutral) may be conveyed by a complex combination of features including *voice timbre* (e.g. a brighter voice is associated with a more positive emotion than a darker voice, possibly because smiling makes the voice timbre brighter) (Laukkanen et al., 1997; Waaramaa et al., 2006).

1.3 Analysing the Prosody of Teacher Talk

From the methodological perspective to understand what happens in classroom, we have to take into account that perceptions of paralinguistic and nonverbal characteristics of talk are, to a large extent, subconscious (see e.g. Zald, 2003). These subconscious characteristics influence discussion processes, as according to Brazil (1978), intonation in discourse illustrates the interaction. Various studies have shown the effect of prosodic aspects on the listeners' opinion of the speaker and perceptions of his/her personality (e.g. Addington, 1968; Lukkarila et al., 2012; Scherer, 1972; Zellner Keller, 2004). Prosodic (or *suprasegmental*) features intensify what is said or add meaning to the segmentals (phonemes). Additionally, several studies have indicated that prosodic features can even reverse the meaning of a message (see e.g. Laver, 1991; Lehiste, 1970; Scherer & Giles, 1979) and some studies have shown that when linguistic-semantic content and prosodic-paralinguistic content are contradictory, the latter usually wins (Lyons, 1977). Therefore, we need a better understanding on what kind of role prosodic aspects play in teacher's talk.

Divergent intonation patterns are also known to be used to praise and encourage students, to minimise the embarrassing effect of a wrong answer and to open and close discussions (Hellermann, 2003). Furthermore, intonation may be used to mark text cohesion (e.g. Halliday & Hasan, 1976), which improves speech comprehension. Studies have also shown that a teacher's dysphonic voice quality (e.g. irregularities in the sound signal, perceived as vocal fry or hoarseness) negatively affects students' comprehension of instruction (see Imhof et al., 2014; Lyberg Åhlander et al., 2014; Rogerson & Dodd, 2005). On the other hand, the classroom activity is also affected by environmental factors, like the size of the student group and classroom acoustics. For instance, a noisy environment requires a higher volume. Consequently, a teacher may unintentionally modify other prosodic aspects, like intonation or voice quality, which may restrict the natural use of prosodic variation in conveying content or even provoke contradictory connotations. This may influence teacher's talk and is another reason to investigate possibilities and limits of prosodic analysis as an approach for gaining insights into classroom talk.

1.4 Aims

The motivation for this article is that the content features of talk might not be sufficient for describing and understanding the true nature of classroom activities. Therefore, methodological development is needed. This article aims to identify novel methods for studying teacher talk, combining both content and prosodic perspectives in this analysis. We concretize the methodological approach in light of two selected physics lessons. Special attention will be paid to the applicability and restrictions of the selected methods for analysing the features of talk. To illuminate the method, we advanced two research questions:

(RQ1): How were the prosodic features of teacher talk influenced by the contextual factors of the authentic classroom, such as noisy classroom conditions?

(RQ2): How did the teacher's use of prosody vary between different kinds of talk patterns?

2. Method

2.1 Context and Data

The present work is an exploratory case study based on data obtained in an authentic classroom setting. Twenty-seven seventh-grade students and a teacher worked in a computer-supported inquiry science classroom during a nine-week physics course (27 hours of teaching and studying). One researcher observed the lessons and took ethnographic field notes during classroom observations (Derry et al., 2010). Based on these observation notes, we selected two lessons that included different types of teacher's talk (see, next section) for future analysis. The two respective 45-minute videos of physics lessons from which both the teacher's and students' dialogue were transcribed, served as data for the present study. The teacher played a central role in planning the practical organisation of the project, and she was not given any specific instructions regarding her role as a teacher in the project. The teacher was fully responsible for implementing the instructional design without interference from the researchers. One video camera and three audio recording systems taped lessons.

2.2 Analysing Teacher's Talk

In this study, we are interested in teacher talk in classrooms as a medium for pedagogy (e.g. Kumpulainen et al., 2010). First, our analysis focused on how the prosodic features of teacher talk were influenced by the contextual factors of an authentic classroom (RQ1). This investigation is necessary because most prosodic research has been conducted in settings in which participants use their 'natural' voice, while teachers may use their voice differently in classroom conditions, as the classroom is a specific condition with rather high levels of noise due to students. Second, we sought whether (and how) a teacher's use of prosody varied between different kinds of talk patterns (educational dialogues and teacher monologues, RQ2). In-depth qualitative analysis and descriptive statistics were used to analyse and interpret the teacher's talk. The identified 110 episodes of educational dialogue (n=42) and teacher monologues (n=68) were analysed. Frequency counts and illustrative qualitative analyses were combined to explore the teacher talk in detail.

The data analysis of classroom talk was grounded in educational dialogues and teacher monologues; it was also adjusted to sociocultural discourse analysis (Mercer & Dawes, 2014; Niemi, 2016). The talk was analysed sequentially, which means that each utterance in a selected sequence is understood and viewed in relation to the previous utterance in the ongoing discussion. According to Linell (1998), analytical descriptions are thus oriented towards the dialectical achievements of the participants. We identified key episodes related to educational dialogues based on triadic dialogue (an I-R-F pattern, Lemke, 1990) and teacher monologues.

As Nassaji and Wells (2000) have noted, however, this basic structure of triadic dialogue can be used for many purposes, particularly because the nature of the feedback the teacher is providing may vary. In our analysis, we focussed on these variations of teacher's feedback. We based this analysis on Sinclair and Coulthard's (1975) follow-up acts – accept/reject, evaluate and comment – however, we further developed the analysis of follow-up moves. We argue that Sinclair and Coulthard's (1975) follow-up moves may not fully account for the influence of feedback variations that are present in the current inquiry-based science classrooms. Teacher talk has radically changed since 1975, especially in the context of inquiry-influenced science classrooms. While the role of the teacher was once that of knowledge provider, classroom talk today is more based on shared discussions in which teachers try to trigger and support their students' knowledge construction activities. For example, whether a teacher accepts or rejects a student's response will make a difference. We thus broadened the analysis of the I-R-F pattern and labelled follow-up moves as cumulative, promotive and disputational I-R-F patterns (see Table 1 for more details) (see also an analysis of student-student collaboration: exploratory, cumulative and disputational talk, Mercer & Wegerif, 1999).

Table 1

The Talk Patterns Used for Coding Transcribed Data

Class	Description (educational dialogues/teacher monologues)
I-R-F patterns	*Cumulative I-R-F pattern* in which the teacher and student(s) respond positively (and cumulatively) to what the student has said to further construct knowledge by accumulation (no critical voices or extreme excitement). Typical features are *repetitions and confirmations*. (Educational dialogues)
	Promotive I-R-F pattern in which the teacher and student(s) engage constructively with each other's talk so statements may be challenged and counter-challenged, but challenges are justified. In addition, *alternative hypotheses are offered by the teacher*. The teacher actively promotes discussions, and *students' opinions are sought*. (Educational dialogues)
	Disputational I-R-F pattern that involves *disagreement and critical voices from the teacher to student(s)* (e.g. teacher is demanding more from the student). Typically, this pattern involves clear critiques and short exchanges consisting of assertions or counter assertions ('Yes, it is', 'No, it's not'). (Educational dialogues)
Group organising pattern	The teacher talks with a student or a few students, typically to get them to work more effectively and concentrate on physics-related issues. (Teacher monologues)
Teacher presentation pattern	The teacher typically gives physics-related information to students. There are few attempts to engage students to take part in the discussion. (Teacher monologues)
Other	Talk that cannot be classified into any of the categories mentioned above. This can include, for example, the teacher's communication with her colleagues. (Not classified – heterogeneous)

To form a meaningful unit of analysis, typically several utterances, both from the teacher and student, needed to be combined. Consequently, there were cases in which sequentially analysed utterances had characteristics of two or more classes from Table 1. In these situations, the coding of units of analysis was based on the teacher's talk. In addition to three educational dialogue patterns, there were episodes of teacher monologues when only/mostly the teacher spoke. During teacher monologues, the teacher, for example, gave physics-related information without dialogic discussion with students. She also organised the groups to get them to work more effectively. These patterns of talk are referred to as *teacher presentations* and *group organising*, respectively. Finally, there was 'other' talk. All these classes with their descriptions are presented in Table 1. The coding was done with software for qualitative data analysis (ATLAS.ti). This coding (six classes, see Table 1) allowed an examination of educational dialogues and teacher monologues, which enabled talk episodes to be identified as patterns for frequency counts. From the coded data, the durations of the units of analysis were determined as measured in seconds. Subsequently, based on the coded frequency counts of teacher talk patterns were selected for the intonation analysis. We did not conduct an intonation analysis for 'other' talk because this group was so heterogeneous in its contexts, for example the teacher discussed with her colleague (without the students hearing it) a student who probably ran away from school. In summary, in

our analysis we first identified talk episodes, such as (1) cumulative, (2) promotive, and (3) disputational I-R-F patterns in the educational dialogues, as well as (4) teacher presentation and (5) group organising in the teacher monologues. After that, we investigated how the prosody related to these patterns of talk was characterised and whether variations could be identified.

The educational dialogues and teacher monologues took place in Finnish, and the researchers translated the excerpts presented in this article. Pseudonyms were used to report the results. Our analysis revealed typical patterns of how the teacher used her voice regarding different classroom situations. We aimed to select representative excerpts of classroom talk. There are several reasons for selecting these specific excerpts. First, in line with how the prosodic features of teacher talk were influenced by contextual factors (RQ1), we selected excerpts that illustrate prosodic challenges that emerged when teachers acted in authentic classroom settings and that may be a useful starting point for future studies. Additionally, to illustrate how the teacher's use of prosody vary between different kinds of talk patterns (RQ2), excerpts were selected in accordance with the theoretical perspective regarding educational dialogues and teacher monologues. Due to space limitations, only some examples of the talk are illustrated in detail. However, we do not claim that the episodes presented here are necessarily typical of the larger sample. Rather, they were chosen in view of our aim to illustrate the method developed and show that prosodic analyses may be an interesting venue for further research. However, to increase reliability, the data excerpts and the analyses have been actively discussed within our research group. Critical comments and joint analysis efforts have contributed to strengthening the validity of the empirical analysis.

2.3 Acoustic Analysis

In the prosodic approach to our data analysis, we focused on pitch and intonation. Additionally, voice quality was addressed, as it may change involuntarily as a consequence of environmental challenges. For readers who are unfamiliar with the measures within prosody research, we briefly explain the terminology in detail in the next section. Readers who are familiar with this terminology can move directly to the section 'Analyses in the present study'.

2.3.1. Introduction to acoustic analysis

Human voice production can be divided into three main parts: power source, vibrator and filter. Airflow from the lungs provides the *power source* for vocal fold vibration. The vocal tract (space from the vocal folds to the lips and nostrils) acts as a *filter*, thereby colouring and amplifying the sound produced by vocal fold vibration. Through articulation, we modify the vocal tract to produce speech. *Speech* consists of linguistic content and paralinguistic cues (prosodic, suprasegmental). *Linguistic content* refers to the words used, and *paralinguistic cues* are the way the words are expressed and what other sounds or modifications are included (laughter, crying, smiling). Paralinguistic content contains prosodic elements. Prosodic features are said to be suprasegmental, as they are properties of speech units larger than the individual segment. It is necessary to distinguish between the personal, background characteristics that belong to an individual's voice (for example, one's habits influencing the pitch range) and the independently variable prosodic features that are used contrastively to communicate meaning (for example, the use of changes in pitch to distinguish questions from statements).

We can alter our vocal production in many ways. The human voice, like sound in general, has three basic characteristics: *pitch* (how high or low the sound is), *loudness* (how soft or loud it is) and *quality* (timbre/colour, i.e. whether the sound is dark or bright, or sounds tense or lax). In addition, sounds have duration (how long or short the signal is). Prosodic characteristics consist of the manipulation of these aspects. Thus, they include the average pitch and loudness used by the speaker, as well as variations in pitch and loudness during sentences. Variations in pitch during a sentence are called *intonation*. Variations of loudness during a word or a sentence are used to stress (highlight) the important parts of the message. The manipulations of temporal aspects in talk include altering the duration of sounds and talk tempo and talk rhythm, consisting of temporal aspects and pausing. Besides these characteristics, we can change our vocal quality: We can speak in a tenser way or a laxer way, or we can use various types of vocal fold vibration: chest voice, falsetto or vocal fry. Furthermore, prosodic aspects are used to complement, intensify or even contradict one's

conversational content. Prosodic aspects also convey (even subconsciously or involuntarily) information about our psychophysiological state, with aspects like mood, emotion and attitude, as well as our physical status (age, gender, health etc.).

Voices can be studied acoustically. The characteristic of sound that is perceived as pitch originates from the fundamental frequency (F0). This, in turn, in human voice production, corresponds to the number of vocal fold vibrations per second. It is measured in Hertz (Hz); one vibration per second is 1 Hz. The faster the vocal folds vibrate, that is the more vibrations they produce per second, the higher a pitch we hear. The average F0 of a male speaking voice is about 120 Hz and a female voice about 200 Hz; however, the frequencies (i.e. the pitch use) may be somewhat language and culture dependent (Pépiot, 2013). Below, in Figure 1, we present an example of the pitch curve, or F0 curve, made from a sentence read by a female speaker: '**To my chagrin** (in Finnish: 'Harmikseni'), I did not find the basket there any longer.': In this example, the mean F0 of the low-pitched Finnish female speaker is 165 Hz, which corresponds to the E note (E3) on a musical scale. The highest peak (F0 maximum) at the beginning of the sentence is 308 Hz (dis1 or DIS4), and the lowest F0 at the end of the sentence is 122 Hz (H or H2). Thus, the pitch variation range in this sentence is approximately 16 semitones (a semitone, i.e. a half step, is the smallest interval in Western tonal music). In general, peaks on an F0 curve are associated with sentence stress (emphasis).
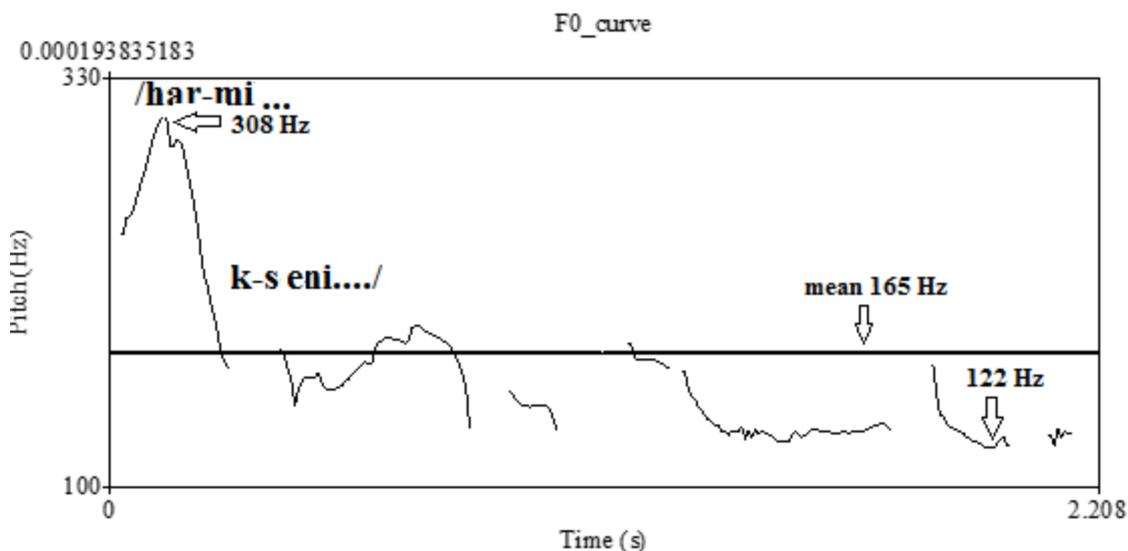

Figure 1. An example of pitch curve, or the F0 curve, and analysis. X-axis: time (in seconds), y-axis: fundamental frequency (F0), pitch in Hertz (Hz, 1 Hz is the inverse of the duration of one vocal fold vibration, thus telling how many vibrations fit in one second of time). The high peak that is clearly seen at the beginning results from sentence stress placed on the word 'Harmikseni' (In English: 'To my chagrin').

The main acoustic correlate of perceived loudness is the *sound pressure level* (SPL), which is most often measured in decibels (dB). Voice quality, in turn, can be acoustically studied, for example, with spectrum analysis. There, the sound is divided into components (harmonics). These components are present simultaneously. How strong these components are in relation to each other affects how the voice sounds, that is, whether the voice's timbre is bright or dark. In a bright voice, there are stronger components in the high-frequency range compared to a darker voice. The overall tilt of the spectrum tells how the voice is produced. The tilt is steeper in a soft or breathy voice (see Figure 2).
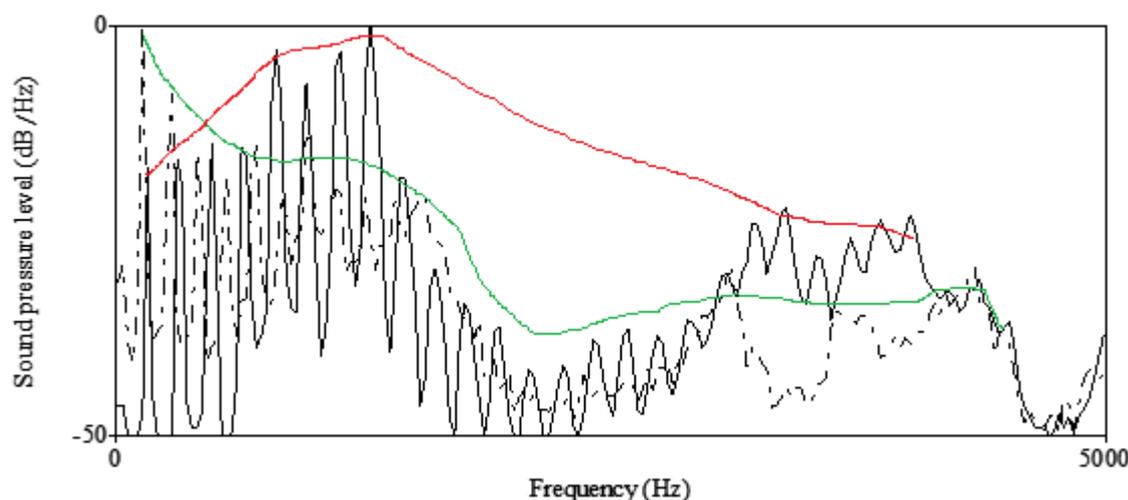
Figure 2. An illustration of two examples of long-term average spectra (LTAS) from two [a:] vowel samples from the same female speaker. The solid line describes LTAS from a pressed voice, while the dotted line shows a breathy voice. The spectrum tilts more steeply in the sample with a breathy voice. This means that the harmonic energy declines faster as a function of frequency. The curves drawn on top of the spectrum (green for breathy and red for pressed) show this phenomenon. The higher the peaks are in the spectrum, the more energy there is in the expression, and the louder the voice sounds.

2.3.2. Analyses in the present study

In the present study, a fundamental frequency (F0) analysis was conducted to give physical correlates of pitch and intonation for sentences that were classified to represent the five teacher-talk patterns mentioned in Table 1: cumulative, promotive, and disputational I-R-F patterns; teacher presentations; and group organising talk. The F0 analysis was performed using Praat software (Boersma & Weenink, 2006, Version 6.0.21). We studied the F0 curve both qualitatively and quantitatively. In the latter approach, we measured the mean, range and standard deviation (SD) of F0. The SD of F0 illustrates F0 variation in intonation more reliably than the F0 range, as the latter can be affected by unintentional voice quality-related matters (like the use of vocal fry with a very low F0) or random errors in the automatic F0 analysis. Voice quality was illustrated through LTAS and Praat analysis.

3. Results

3.1 Prosodic Challenges When Teachers Act in Authentic Classroom Settings (RQ1)

Our methodological approach offers possibilities to show prosodic challenges that emerged when teachers acted in authentic classroom setting. We found that when the teacher interacted with students in authentic (meaning often rather noisy) classroom conditions, she used a loud voice, which led to a heightened pitch. Figure 3 illustrates this phenomenon in terms of an F0 curve. As we can see in Figure 3, there is a part of teacher presentation, interrupted by a question from a student, followed by the teacher's answer to the question.

Teacher: Speaking of next week's exam that we thought about today…

Student: **So, when it will be next week?** (In Finnish: 'Eli millon se on ens' viikolla?')

Teacher: On Wednesday of next week.

Here, the teacher's mean F0 is 296 Hz (d1) in the beginning, which represents her general way of speaking to the whole class. At the end part of the curve, the mean F0 is 220 Hz (a), as she answers a student's question seemingly using her natural conversational volume. Thus, at the beginning she raised her pitch by circa 5 semitones to speak to all the students. This may also restrict the habitual livelier use of pitch variation, as can be seen by comparing the beginning part of the F0 curve in Figure 3 with the F0 curves seen in the other figures.

In Figure 4, we illustrate this exchange in terms of voice quality. The black line shows a spectrum of the part: 'Speaking of next week's exam that we thought about today…' In this part, the teacher speaks loudly, and her perceptual vocal quality is pressed (see e.g. Kankare et al., 2012; Waaramaa & Kankare, 2013; Waaramaa et al., 2014). The grey line shows the spectrum of the part: 'On Wednesday of next week.' In this part, the teacher's voice sounds more relaxed. Thus, in pressed phonation, there is more sound energy at the higher frequency range than in the ordinary phonation type. The use of a pressed vocal quality poses more biomechanical load on the vocal folds than the use of an ordinary, relaxed voice. Additionally, this example illustrates that a teacher seemed unintentionally modify prosodic aspects, which may restrict the natural use of prosodic variation in conveying content or even provoke contradictory connotations. For instance, a teacher may involuntarily sound angry when trying to get his/her voice heard over background noise. This, in turn, may influence how teacher's talk is interpreted in classroom.
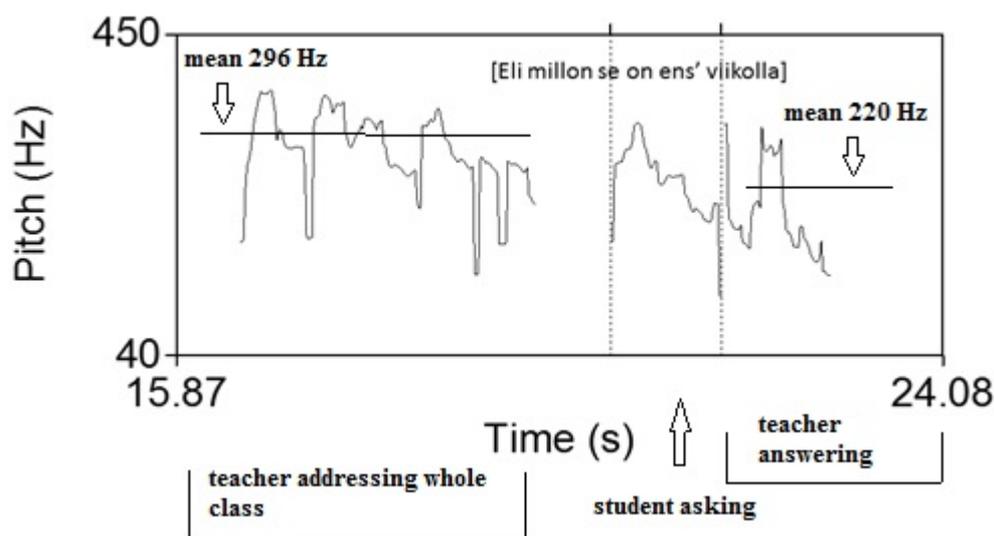

Figure 3. An example of how the teacher raises her pitch by circa 5 semitones (left part of this figure) when raising her voice to speak to the whole group of students (compared to when she is speaking to a single student, right part of the figure).
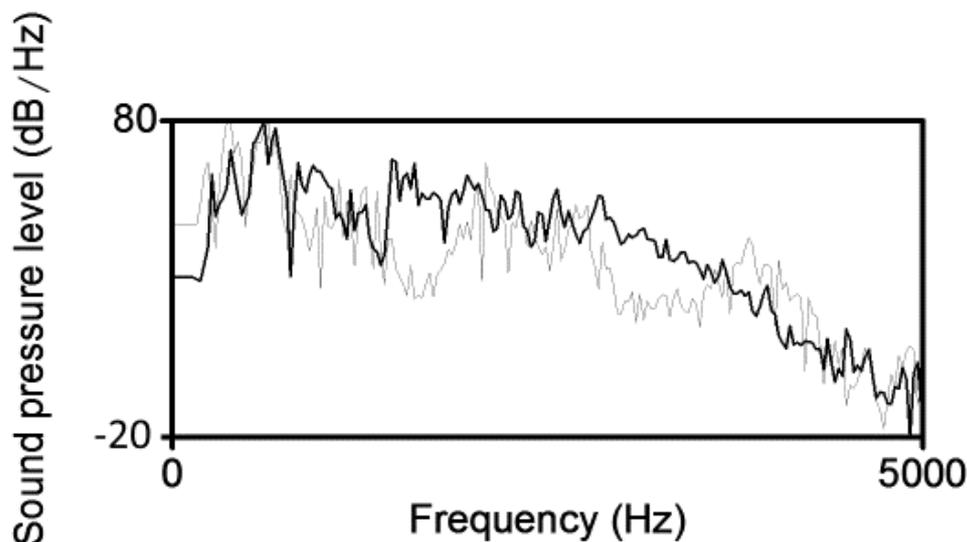
Figure 4. Two long-term average spectra of teacher talk. Y-axis: mean sound energy in dB, x-axis: frequency in Hz. The black line represents a pressed and loud voice (addressing the complete group of students), while the grey line represents a relaxed voice (talking to an individual student).

3.2 Differences in a Teacher's Use of Prosody (RQ2)

In Table 2, we summarise the F0 characteristics in the five patterns of talk studied (see Table 1 for a description of the patterns). Even though based on this small sample size, it is impossible to claim direct correspondences between the patterns of talk and intonation patterns, some typical features could be identified based on this exploratory case study. In general, cumulative I-R-F patterns seemed to use a moderate mean pitch level and a moderate pitch variation (see Table 2), with frequent word stresses that were realised using the same pitch pattern (see Figure 5). Promotive I-R-F patterns showed a low mean pitch, a wide pitch range and large emphatic sentence stresses (high peaks in the F0 curve; see Figure 6). Disputational I-R-F patterns seemed to use moderate mean pitch level with large pitch variation and strong emphatic stresses (see Figures 7 and 8). Teacher presentation resulted in a high mean pitch level with small pitch variation (Figure 9), while group organising was characterised by a relatively high mean pitch level with moderate pitch variation (Figure 10). In the following sub-sections, our methodological approach is exemplified with empirical examples. We demonstrate how language was manifested in various talk patterns and what kinds of prosodic features were typical for each type of talk pattern. For each type of talk listed in Tables 1 and 2, we first introduce a representative example of a talk episode, followed by a representative figure (graph of a prosodic phase) illustrating how the teacher's intonation is used and varies. Within the selected representative sentences, bold words refer to stressed words in the sentence.

Table 2.

Teacher's F0 (Pitch) Variation and Range in Hz and in Semitones.

Teacher's talk	Mean F0 (Hz)	F0 range (Hz)	SD (Hz)	SD (semitones)
Cumulative I-R-F - patterns	232	112–335	46	7
Promotive I-R-F - patterns	199	98–310	55	9.8
Disputational I-R-F - patterns	230	84–346	66	10
Teacher presentation	290	193–374	37	4.4
Group organising	248	94–419 56-419	55	8

3.2.1. Teacher's talk prosody in educational dialogues

Cumulative I-R-F patterns were rare and emerged nine times (8.2%) with a total duration of 333 seconds (12.2%). They involved speakers in pleasant, uncritical exchanges that built towards a common understanding through accumulated repetition and confirmation. From a prosodic perspective, cumulative I-R-F patterns were associated with a relatively narrow pitch range and frequent word stresses that were realised using the same pitch pattern. Typically, the intonation pattern repeated itself, without extreme deviations from the mean pitch level.

The following excerpt 1 of cumulative I-R-F pattern shows how the teacher cumulated dialogue. First, she conformed: 'Well, now, it's here', and she asked what material a jar of jam is made from. The student Pekka responded. This was followed by a new question from the teacher and another response from Pekka. Subsequently, the teacher inquired what happens to materials when they are heated, and she accepted a trivial answer from Joel, a student: 'The material gets warmer' (i.e. constructing positively). Then, the teacher tries to remind the students of a video in which the same phenomenon was shown. In this way, the common knowledge is constructed further. The teacher repeats her student Elvira's answer that, when matter becomes warmer, it expands (conformation and repetition).

Figure 5 is selected from excerpt 1 to show a typical example of the flow of the cumulative I-R-F pattern. In this example, the mean F0 is 232 Hz, and the SD is 46 Hz, that is, seven semitones, and the total F0 range is 112–335 Hz. Furthermore, the graphic lines show that the intonation pattern repeats itself, without extreme deviations from the mean pitch level. Between the vertical lines is a sentence: '**Expands** (in Finnish: 'laajenee'), when matter becomes **warmer** (in Finnish: 'lämpenee'), it **expands** (in Finnish: 'laajenee'). Now, when the jar is made of glass, the lid is made of metal. Does it expand in the same way?'

Excerpt 1: An example of cumulative I-R-F patterns

Teacher: Well, now, it's here. Think about a jar; what material is a jar usually made of? A jar of jam.

Pekka: Glass.

Teacher: And what material is the lid of the jar, typically?

Pekka: Metal.

Teacher: Think about glass and metal; when you put the jar under hot water, what happens to it? What happens to matter when it becomes warmer? When getting warmer, matter, what…?

Joel: It becomes warmer.

Teacher: Becomes warmer, but at the same time …? Quite in the beginning we made those, there was … You have the kind of a video there, with the hole and the bullet, and the bullet is heated.

Elvira: Expands.

Teacher: Expands, when matter becomes warmer, it expands. Now, when the jar is made of glass, the lid is made of metal. Does it expand in the same way?

Liisa: The lid gets larger.

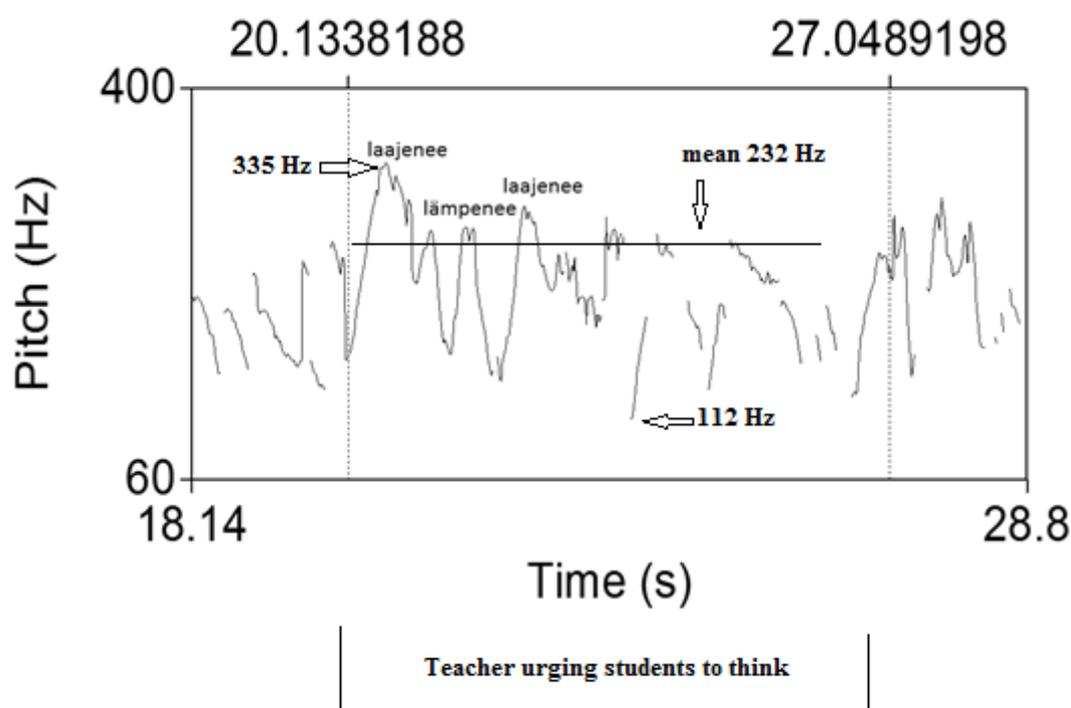Teacher: Yes, the lid gets larger than the jar, so then we get it open.

Figure 5. An example of a cumulative I-R-F pattern.

During *promotive I-R-F patterns*, the teacher engaged constructively with students' ideas, trying to trigger productive collaboration (to trigger exploratory talk, see, Mercer & Wegerif, 1999). Promotive I-R-F patterns were applied fairly actively (*n*=20, 18.2%; total duration of 11.4 minutes, 25%). The prosodic analysis illuminated that promotive I-R-F patterns were associated with a wider pitch range than cumulative I-R-F patterns and larger emphatic sentence stresses than the cumulative I-R-F patterns presented previously.

The following excerpt 2 is a typical example of a discussion between the teacher and her students. The teacher asks her students to think of the three materials and infer which expands differently from the others. The teacher urges her students to consider a phenomenon and converse about it, and she offers an immediate verbal response to her students' reactions. As we can see, the teacher engages students to actively encounter the phenomenon: 'Now, think of these three materials. Can you infer which one of these expands differently from the others?' When Juuso's response is correct, the teacher becomes excited and immediately gives him positive feedback: 'Correct! Well done!' The teacher continues to explore the issue and asks her students to make a hypothesis regarding what happens to metals when they are heated. In this case, the teacher tries to engage her students to consider and justify their answers (aiming to promote students' exploratory talk). She

also offers alternative hypotheses: 'Well, what happens to metals when they exp … become warmer? Do they shrink, stretch or not change?' After Ilkka replies correctly, the teacher again provides positive feedback.

Figure 6 highlights the discussion between the teacher and a student. For Figure 6, the mean F0 is 199 Hz, and SD 55Hz, i.e. 9.8 semitones, and the total range is 98–310 Hz. We can see that promotive I-R-F patterns showed a wider pitch range and larger emphatic sentence stresses than the cumulative I-R-F patterns presented previously. Furthermore, an emotional state of excitement involves a high arousal level that is seen in high F0 peaks in the intonation curve: '**Correct! Well**…' (in Finnish: 'Aivan! Hyvin…') The two high peaks in the pitch curve at the end of the sample reflect stressed words, expressing the teacher's excitement when getting a correct answer from her student: 'Correct! Well done!' (in Finnish: 'Aivan! Hyvin pää(telty)!') The last syllables in the parentheses are expressed in a whisper. The excerpt also shows how the intonation curve drops in a question in Finnish (the first half of the picture).

Excerpt 2: An example of promotive I-R-F patterns

Teacher: If you think of these three materials, could you infer which one of those expands differently from the others?
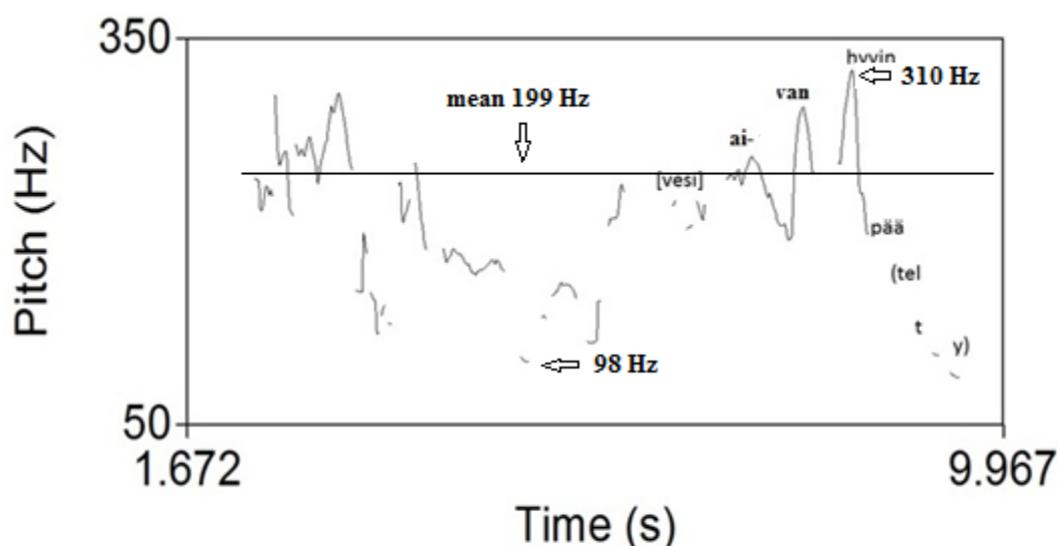
Juuso: Well, water.

Teacher: Correct! Well concluded!

Juuso: Do I win a prize?

Teacher: Nope. Well, what happens to metals when they exp … become warmer? Do they shrink, stretch or not change?

Ilkka: They stretch.

Teacher: Good! The first page is completed.

Figure 6. An example of promotive I-R-F -pattern.

Finally, *disputational I-R-F patterns* were characterised by disagreements, by teacher disagreeing and showing critical approach to student response(s) and by short, often confrontational, interchanges from the students. It emerged 13 times (11.8%) with a total duration of 346 seconds (12.7%). Figures 7 and 8 below illustrate that during disputational I-R-F -patterns, the pitch variation seemed to be the greatest and sentence stresses strongest.

In the following excerpt 3, we can see how the teacher gets frustrated when trying to motivate the students to think about the insulation properties of a thermos bottle. The teacher asks many times if the students could explain why the inner surface of the thermos bottle is glossy without giving them sufficient resources. There is also evidence that the students are not listening actively to the teacher, as Markus asks regarding the glossy surface: 'On the inside, you mean?' even though the teacher has consistently been talking about the inner surface of the thermos bottle. Finally, when Markus responds to the question, the teacher disagrees with him: 'That isn't enough, Markus, that it keeps the drink warm.*'* In practice, the teacher demands more information from Markus.

Of the interactive patterns, pitch variation seems to be largest and sentence stresses strongest for disputational I-R-F -patterns. For Figure 7, the mean F0 is 230 Hz, and SD 66 Hz, i.e. 10 semitones, and the total range is 84–346 Hz. In Figure 7, we can also notice that the sentence stress is on the word 'enough' (in Finnish: 'riitä'), which is indicated by the highest peak in the intonation curve: '**That isn't enough** (in Finnish: 'Toi ei riitä...'), Markus, that it keeps the drink warm.' When comparing Figure 7 to Figure 6, high F0 peaks can be seen in both figures reflecting a high arousal level in the teacher's talk. However, in Figure 6, a positive emotion of excitement was expressed and in Figure 7 a negative emotion, perhaps frustration considering the content of the teacher's talk. Thus, the emotional valence (positive, negative or neutral) cannot be directly concluded from the acoustic cues of the arousal level e.g. related to intonation (Bänziger & Scherer 2005).

Figure 8 below illuminates how this disputational I-R-F pattern continues with similar significant prosody variation. Here, the teacher is still not happy with the students' study process, and she is still demanding more from the students. In the episode represented in Figure 8, the teacher's mean F0 was 210 Hz, the SD was 46 Hz, i.e. ca 8 semitones, and the range was 122–317 Hz. Thus, in the continuation of disputational I-R-F pattern the pitch variation continues to be great and sentence stresses strong: 'It is kind of a **fact** (in Finnish: 'fakta') why a vacuum bottle **is used** (in Finnish: 'käytetään')**,** but you should now explain why the **glossy** (in Finnish: 'kiiltävä') interior is helpful there.' The very low F0 values in the total range reflect the use of vocal fry phonation, especially in sentence endings.

Excerpt 3: An example of disputational I-R-F -patterns

Jan: A thermos bottle is a bottle that is heat insulated from its environment as much as possible.

Teacher: Yes. Does it explain why the inner surface is made glossy? It is one of the many ways by which it is made a good insulation, the structure of the whole bottle, but…

(off-topic discussions between the students)

Teacher: YES, but why does the glossy surface keep it warm unlike a red surface, for example?

Jan: I don't know.

Teacher: And you cannot find any explanation, can you, if you go and study the material there?

[miscellaneous noise]

Teacher: Hey – now! You were supposed to think now of the glossy inner surface of a thermos bottle. What might be [I know] the reason? Well?

Markus: On the inside, you mean?

Teacher: Uhum, on the inside, yes. Have you ever looked into a thermos bottle? [Miscellaneous noise, The teacher's comments to one student.] Think; search for information. It can be found there in the heat transfer mechanisms section.

Teacher: You can't search for anything, can you?

Markus: Me? Yes, watch out; it's coming soon…keeps warm….

Teacher: That isn't enough, Markus, that it keeps the drink warm. It's like a fact why people use a thermos bottle, but you should now explain why the glossy surface helps there. So, because what...?

Markus: I don't know.

Teacher: Well, it cannot be necessarily found from Wikipedia now.

[miscellaneous noise]

Teacher: It is kind of a fact why a vacuum bottle is used, but you should now explain why the glossy interior is helpful there.
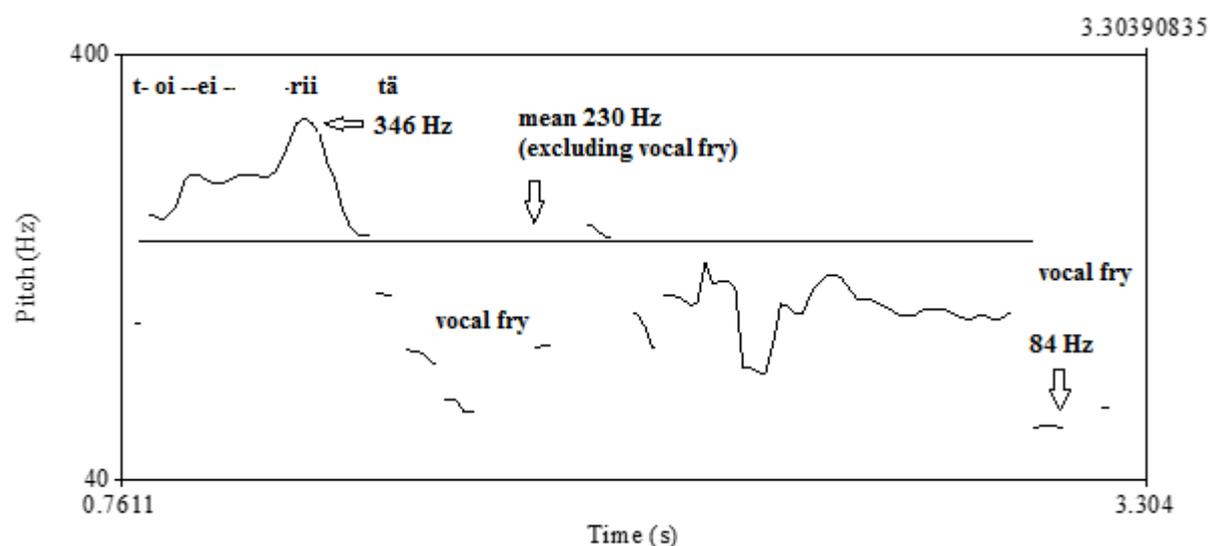

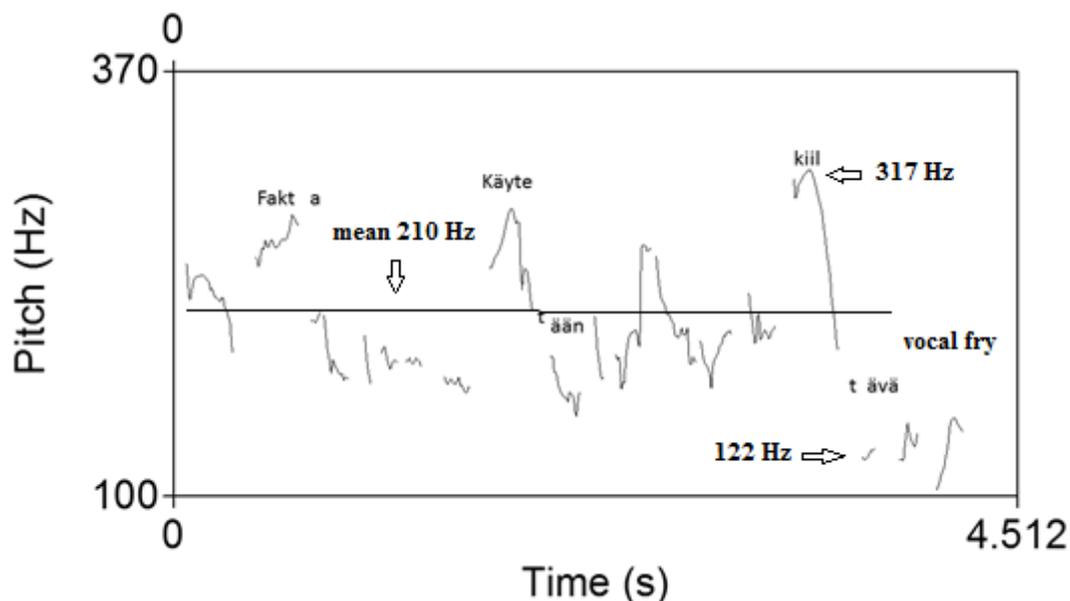*Figure 7.* An example of disputational I-R-F -pattern

Figure 8. Another example of disputational I-R-F -pattern, a continuation from Figure 7.

3.2.2. Teacher's talk prosody in teacher monologues

The teacher's talk initiated by a student's question or her talk at the beginning of different sections of the lesson was labelled as *teacher presentation*. The teacher gave general instructions to the students, so they could start working on their tasks, the teacher interrupted the students' working so they could begin reviewing the correct answers to the problems, or the teacher gave a short lecture about the theory when students faced challenges while solving problem; these are typical examples of when talk classified as teacher presentation emerged. In general, it can be said that when there was a need for straightforward instruction (Mercer, 1995) the teacher's talk had characteristics of teacher presentation. Overall, teacher presentation was a common type of talk, comprising one fourth of the total duration of different talk patterns (n=27, 24.5%; total duration 11.5 min, 25.3 %).

As an example, in the following excerpt 4 the teacher does not provide physics-related information, but rather general instruction to the whole group about the lesson plan of the day. She also gives general feedback to the students about combining information from different sources when taking their previous exam. There are no attempts (e.g. questions) to invite students to take part in this discussion, so the conversation can be referred to as teacher monologue. Figure 9 displays an example of teacher presentation talk: 'There were some **minor difficulties in the answers** (in Finnish: 'pieniä ongelmia vastauksessa') to the exam **last week** (in Finnish: 'viime viikon'). You were not **able to combine** (in Finnish: 'osannu yhdistellä') pieces of information …' The mean pitch of the teacher's talk is relatively high (290Hz), and there are no wide changes in F0 during intonation. The range of F0 variation comprises frequencies from 193 to 374 Hz, and the SD is 37 Hz, i.e. about 4.4 semitones.

Excerpt 4: An example of teacher presentation

Teacher: There were some minor difficulties in the answers to the exam last week. You were not able to combine pieces of information (or) find it, so I think that we will practice a little for the exam. There are similar types of questions to those that will be on the exam, so let's take a look (at these). First, go through (the problems) yourself or with a pair or a group, and think about how you would answer. Together, try to find what kind of answer would be good when you have to combine pieces of information from many resources now.
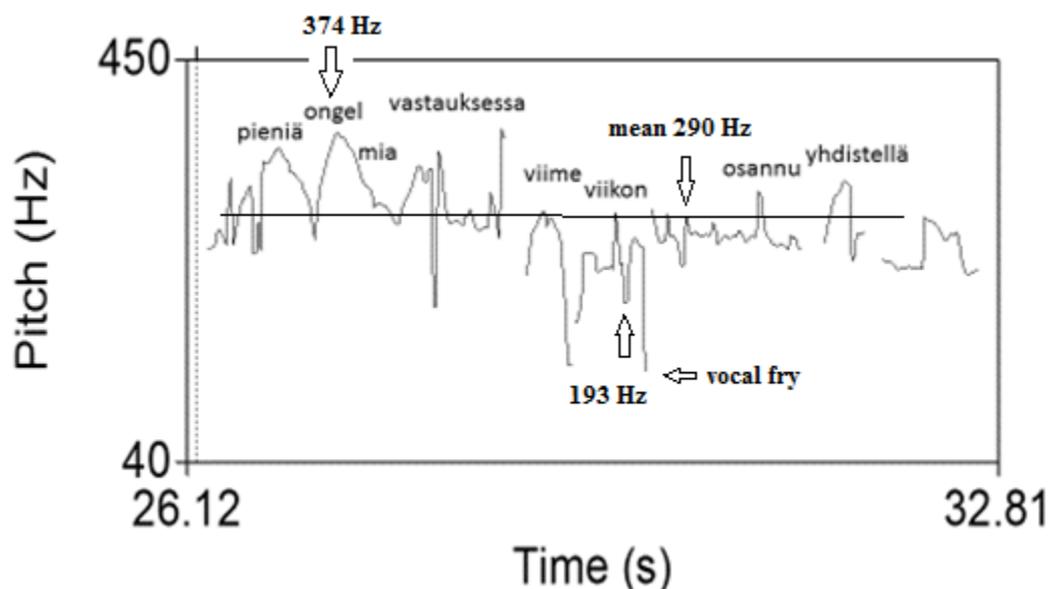
Figure 9. An example of teacher presentation talk.

In addition to teacher presentation, there were 31 units of analysis (28.2 %) belonging to *group organising*, but the total duration (9.0 min, 19.8 %) of those units was not that high, mainly because excerpts were usually short remarks and comments from the teacher to the students relating to their behaviour and studying methods. Group organising, like teacher presentation, was characterised by teacher monologue. In general, there was a need for group organising at regular intervals when students solved the problems themselves (6.5 min, 25.4 %, cf. teacher presentation: 3.9 min, 15.4 %). When the section of the lesson was more teacher-led by nature (the students started to go through the correct answers with the teacher), the teacher had to guide the students less frequently to concentrate on the teaching (2.5 min, 12.5 %, cf. teacher presentation: 7.6 min, 38.0 %).

In the following excerpt 5, the teacher explains to the students that the problems they are going to solve are a rehearsal for the exam. When a student points out that he did not get a problem sheet, the teacher says, with a twinkle in her eyes, that some of the students may have taken more than one (same) problem sheet, as there were not enough papers in the stack. Even though there are utterances both from the teacher and the student, there is no typical triadic dialogue visible, and the conversation can be referred to as teacher monologue without true collaboration between the participants. The main motive of the teacher was to get the problem sheets for everyone so the students can start revising physics-related issues for the exam.

A situation given below illustrates the use of prosody during group organisation. In Figure 10, we can see how the teacher says that the exercise 'is a **rehearsal** (in Finnish: 'harjoittelua') for the exam', and a student responds that '**I didn't get one'** (in Finnish: 'Mä en saanu.**')** Then, the teacher answers that 'There should be (more) **in the stack** (in Finnish: 'Siin pinos'), so perhaps somebody took more than one. See, **the most hard-working** (in Finnish: 'ahkerimmat') students do two (papers).' As we can see from Figure 10, the teacher's mean F0 was 248 Hz, the absolute minimum 56 Hz representing vocal fry (creaky sound), 175 Hz was the lowest F0 without vocal fry, and the maximum was 419 Hz. Thus, the total frequency range was 15 semitones, i.e. 1 ¼ octaves. The SD of F0, which reflects the intonation range more reliably, was 55 Hz, i.e. approximately 8 semitones.

Excerpt 5: An example of group organising

Teacher: This is a rehearsal for the exam.

Jenna: I didn't get (one).

Teacher: There should be (more) in the stack, so perhaps somebody took more than one. See, the most hard-working students do two (papers).
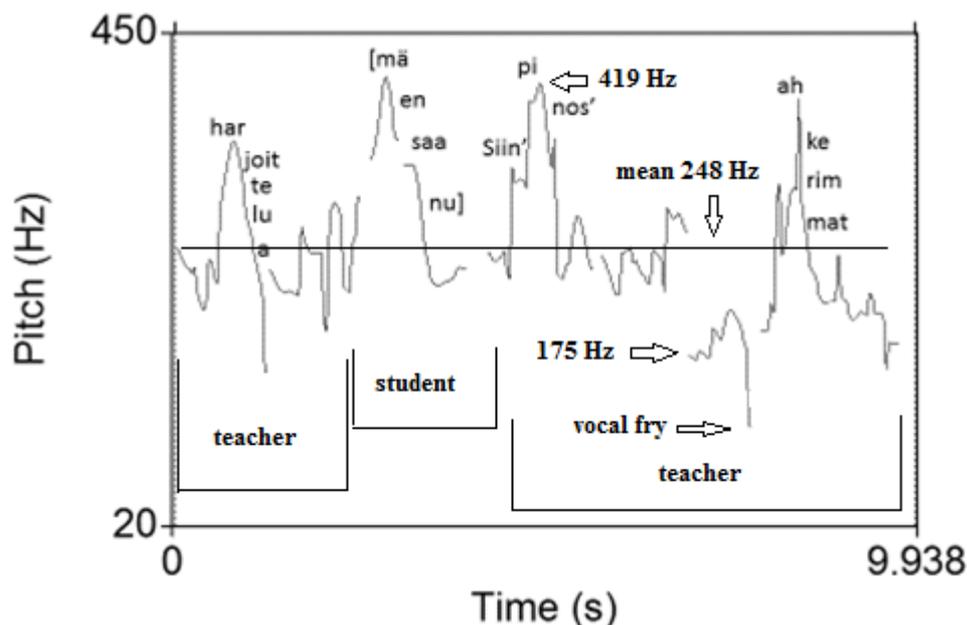

Figure 10. An example of group organising.

4. Discussion

The present study is a first step towards developing a new method of analysing classroom interaction. We investigate the potential of prosodic analyses of teacher talk. Our argument is that analysis of prosodic features has been underrepresented when classroom talk is analysed. We claim that in addition to analysing the content of talk (focusing on *what* is said), analysing the prosodic features of talk (focusing on *how* something is said, thus considering elements as intonation, volume, and pace) is also important. In some cases, it might be as important as – or even be more important than –what is actually said. For example, a teacher asking a student, 'What do you think about this?' may be simply inquiring for a student's opinion. However, the same question – in exactly the same wording – can be conveyed in such way (by changing intonation and stressing other words) that a student really feels involved in the discussion process and appreciates being invited to share his/her ideas. At the same time, exactly the same words can be pronounced in such a way that the student feels threatened and reprimanded for not listening attentively.

In this article, the methodological development grounds on a notion that voices can be studied acoustically. We focused on teacher's talk in an authentic classroom, and two research questions were addressed in relation to the general aim of investigating the potential of prosodic analysis. Knowing that a classroom is far from a laboratory setting, we investigated how the prosodic features of teacher talk were influenced by the contextual factors of the authentic classroom (i.e. an often quite noisy environment) (RQ1). With our methodological approach, we were able to illustrate some specific prosodic challenges. Our findings showed that when the teacher acted in the authentic classroom setting, she often used her voice in a different way. The results show that when addressing the complete classroom, her voice was more raised, resulting in a more pressed voice (indicated by a higher pitch) than in other occasions, such as talking to the student in a one-to-one way or to small group of students when guiding them. In the latter situation, the voice was more relaxed and thus closer to her natural voice. We argue that how teachers use their voice may have an influence

on teachers' health, teacher-student interaction, and classroom climate. Firstly, the risk of vocal fatigue increases when using a pressed voice (e.g. Kankare et al., 2012). Secondly, a pressed voice quality is related to the expression of anger (e.g. Laukkanen et al., 1997; Waaramaa et al., 2010; 2014). Therefore, speaking in a large and noisy classroom may lead to involuntary and misleading prosodic characteristics. These characteristics can be interpreted as shouting in anger, which may affect negatively on teacher-student interaction and classroom climate (see, 3.1). This may be disconcerting as a high-quality teacher–student interaction and a supportive classroom climate is one possible protective factor against the negative impacts of learning (Kiuru et. al., 2012).

The second research question focused on studying how the teacher's use of prosody varies between different kinds of talk patterns (RQ2). Therefore, we first identified talk episodes, such as (1) cumulative, (2) promotive, and (3) disputational I-R-F patterns in the educational dialogues, and (4) teacher presentation and (5) group organising in the teacher monologues. Next, we checked how the prosody related to these patterns of talk was characterised and whether differences could be identified. We found that *cumulative* I-R-F patterns seemed to use less pitch variation and word stress patterns were often repeated here. On the opposite, a wide pitch range and clear emphatic sentence stresses with large F0 jumps characterised *disputational* as well as *promotive* I-R-F patterns. Thus, the intonation pattern in cumulative I-R-F patterns seems to reflect continuation, while the strong emphatic stresses with relatively wide pitch intervals marks a contrast e.g. between the student's answer and the teacher's instruction. In promotive I-R-F patterns, some strong accents with high pitch peaks were used to acknowledge correct answers and to give support. Regarding teacher monologues, *teacher presentation* used a high mean pitch, a narrower pitch variation and a more pressed voice quality, while *group organising* was characterised by a relatively high mean pitch level with moderate pitch variation. In sum, by combining the prosodic and content characteristics of teacher's talk, we were able to identify initial variations in how the teacher used her voice in diverse educational dialogues and teacher monologues.

4.1 Limitations and critical issues

The strength of this study is that, along with studying the content of the talk, it pays attention to the potential offered by the prosodic perspective of teacher's talk that has rarely been explored to date. However, there are several limitations and critical issues to consider as this study was an initial attempt to illustrate how the teacher's intonation varies depending on the situation. First, this study is exploratory in nature, and although we were able to show that different prosodic characteristics are somehow related to distinctive patterns of talk content-wise, additional explorative and hypothesis-testing research is needed to analyse this relationship more specifically. Second, as this case study - like case studies in general - is based on a small sample, all limitations thereof should be duly considered.

Moreover, there are three more limitations to our study that are related to the use of prosodic analysis in general in this type of research settings. The third limitation concerns the use of acoustic speech methodology in an authentic classroom setting and comprises three problematic aspects: (1) Although the technology is available, it is not necessarily easy to get the hardware needed (especially when using it on a larger scale) and to use this hardware to capture voices in classrooms without compromising the authenticity of the setting; (2) the teachers may tend to use their voice in different ways depending on the specific conditions within the classroom; and (3) authentic classroom conditions may hamper the quality of audio recordings and thus limit the usability of the method.

The fourth and fifth limitations also pertain to acoustic speech research in general, both of which may make it more difficult to establish a catalogue of normative data on classroom interactions. The fourth limitation is that the interpretation of how voice is used might be culturally bound (Waaramaa, 2014; Waaramaa & Leisiö, 2013), and this aspect was not considered in this study. On a general level, this limitation might also make it more difficult to compare findings from different classrooms around the world. The fifth limitation is that language specificity might form another barrier for the comparability of the research findings. Specifically, when studying collaboration, language is always the central aspect under investigation, and with regard to prosodic analysis, specific features of different languages may have specific characteristics (see,

Method section). Therefore, we briefly discuss the specifics of intonation patterns and features of Finnish language (compared to other languages) in the remainder of this paragraph. In Finnish, sentence stress and intonation do not serve linguistic purposes to the same extent as e.g. in Swedish or English, as Finnish takes advantage of enclitics. The intonation curve is typically declining in statements, and a relatively smooth and high intonation pattern is used to express continuation (Aaltonen & Wiik, 1979). A pitch rise in sentence endings has been regarded as untypical for Finnish language, even though lately it has become a characteristic of teenagers' talk (Routarinne, 2003 a and b; Härkönen, 2016). In general, Finnish talk has been described as characterised by soft phonation, a low mean pitch, small intervals in intonation and a relatively 'tame' expression of emotions (Hakulinen, 1979). On the other hand, despite the differences between the languages, prosody has been analysed in other fields, e.g. the therapist–patient dialogues (e.g. Leszcz, 2017). From this research, we know that differences between the languages can be considered and dealt with. In the present study, we investigated the intonation pattern of a teacher's talk and our findings can be considered to be in line with earlier results e.g. reported by O'Connor and Arnold (1973) for English language. However, even though we can take differences between languages into account, it could be interesting to explore the value of prosodic analysis in view of analysing teacher talk in different languages.

4.2 Directions for future research

In this section we put forward many opportunities that this new method may bring, and we relate this to some elements to be further developed and explored in future research. First of all, we see the potential for developing our methodological approach towards (semi-)automatic analysis of audio and video data. In a first phase, a possible methodological application could be to identify interesting discussion phases based on the prosodic features, which can then be further analysed and interpreted by educational researchers. Based on the results of this explorative study, we are optimistic that using prosodic analysis in such a semi-automatic way is a likable future application, and even if it means that the identified phases still need to be interpreted by researchers, this is a promising venue, as often researchers have an enormous amount of data, and thus being able to use prosodic analyses to pre-process and reduce this amount of data for manual coding would be a useful application. In a second phase, future research could focus on investigating whether it is possible to move to fully automatic analyses of talk, based on prosodic analyses.

A second opportunity and direction for further research is to broaden the scope to also analyse students' talk. While our study focused on teacher's talk, we suggest that next methodological step should be taken by combining prosodic and content analysis to study students' talk, and more specifically student–student dialogues that are happening as a part of collaborative learning. In this respect, prosodic analyses could be applied to identify different types of talk or collaboration on the one hand, while on the other hand it could be used to capture students' (and also teachers') emotions. Earlier research in the field of therapist–patient dialogues has shown that in addition to verbal, non-verbal, para-verbal, implicit and explicit communication, prosodic analyses are useful in capturing and predicting emotions (Leszcz, 2017). The role of emotions has often been underestimated and could be of great importance (Isohätälä et al., 2017). What is particularly important is the consideration of how, when and why students' emotions arise and how they shape interaction (student-student/teacher-student) and affect students' dedication towards collaboration and learning. This may be associated what and how is said in the classroom context (e.g. our results about teacher involuntary sounding angry). Recently, positive activating emotions have been shown to be related to good academic success (Postareff et al., 2017), so being able to capture and analyse emotions through prosodic analyses – and in a next phase do this ad hoc, on the fly and provide teachers with this information through a learning analytics powered dashboard – could be very interesting and valuable future application.

Related to this, a third opportunity that we put forward is the development of new digital tools to support teachers (see also, Harteis, 2018). Automatic prosodic analyses could be an interesting feature to inform teachers of students' collaborative discussions, e.g. by signalling group processes to the teachers. If students' voices could be interpreted on the fly, data from these analyses could be used to create process indicators in an automatic way. By adding information from automated prosodic analysis, existing tools could be extended. As an example, we can think of how a lantern device (Dillenbourg et al., 2011) could be fed by prosodic data. The lantern device of Dillenbourg and colleagues (2011) is a small device with LEDs that is

controlled by students to allow them to indicate which exercises or phase of a collaborative activity they are working on (i.e. by changing the colour of the lamp), and if they have questions for the instructor, it allows them to signal this to the instructor (i.e. by making the lantern blink). The blinking rate is furthermore increased over time, allowing the instructors to see how long students have been waiting for them (for more details, we refer to Dillenbourg et al., 2011). The goal of the devices was to provide the instructors with some awareness of the teams' behaviour. In their implementation, students controlled the tool themselves, but in future extensions, based on automated on-the-fly analyses of the prosodic features of students' collaborative discussions, the tool could provide additional useful information about collaboration processes for instructors.

Finally, a fourth opportunity is allied to teacher training and teachers' professional development. Being able to capture, interpret and understand students' emotions on the fly while engaged in technology-enhanced collaborative learning may be helpful for teachers' professional development. This is also related to the question of how emotional valence — whether positive, neutral, or negative — can be derived from the teacher's talk. Typically, vocal emotions are studied first for their arousal level and second for their valence. In the present investigation, we concentrated on arousal level, displayed by intonation curves. In our future research, we will scrutinise teacher's vocal expression of valence, how the teacher uses his/her voice to convey emotions related to the content of the talk, e.g. when encouraging the students, when expressing contentment or disappointment, and how valence expressed is associated with teacher's talk. In this respect, research needs to focus on triangulating data resources. So far, there is research available focusing on physiological measures of emotions (e.g. with the smart rings, e.g. http://www.moodmetric.com) or heart rate variability measures (see e.g. https://www.firstbeat.com/en/) and self-report measures of emotions (see Oksanen & Hämäläinen 2010; Castellar et al., 2014). We argue that an application could be to add prosodic analyses as a method in combination with these methods, as another source to triangulate from.

To conclude, there is a current trend of exploring more advanced methods to capture social, cognitive, and emotional features of classroom talk, as these novel approaches are needed to meet the analytical challenges of making sense of the processes of learning and instruction (Damsa & Ludvigsen, 2016). The present exploratory study can in this view be seen as one contribution. We showed that acting in an authentic classroom setting might trigger specific prosodic aspects in the teacher's talk. Additionally, we were able to identify differences in how the teacher used her voice and relate those to diverse educational talk patterns. We believe that prosodic analyses may be one novel approach that allows us to understand learning and instruction processes better.

Keypoints

- Multiple methods and techniques are required to understand what happens in classrooms
- Prosodic aspects (features of speech such as intonation, volume, and pace) of talk are under-represented in the field of the learning sciences
- We introduce prosodic analyses as a method to study teacher talk in classroom
- We showed that the teachers' prosody varied depending on different patterns of talk that were identified based on the content.
- This article shows that prosodic analyses may have an added value for research on learning and professional development

Acknowledgments

References

Aaltonen, O., & Wiik, K. (1979). (1979). Suomen jatkuvuuden intonaatiosta. In P. Hurme. (Eds.)*Jyväskylän yliopiston suomen kielen ja viestinnän laitoksen julkaisuja, 18 1. Fonetiikan Päivät (the First Finnish Phonetics Symposium),* (pp. 23-33).

Alexander, R. J. (2001). *Culture and pedagogy: International comparisons in primary education* (pp. 391-528). Oxford: Blackwell.

Addington, D. W. (1968). The relationship of selected vocal characteristics to personality perception. *Speech Monographs, 35*(4), 492-503.

Berry, M. (1981). Systemic linguistics and discourse analysis: A multi-layered approach to exchange structure. *Studies in discourse analysis*, *1*, 20-145.

Bänziger, T., & Scherer, K. R. (2005). The role of intonation in emotional expressions. *Speech Communication, 46*(3), 252-267.

Boersma, P., & Weenink, D. (2006). Praat: Doing phonetics by computer.

Brazil, D. C. (1978). *Discourse intonation II, discourse analysis monographs II* (1st ed.). Birmingham: University of Birmingham, English Language Research.

Castellar, E. N., Oksanen, K., & Van Looy, J. (2014). (2014). Assessing game experience: Heart rate variability, in-game behavior and self-report measures. In Anonymous (Eds.)*2014 Sixth International Workshop on Quality of Multimedia Experience (QoMEX)* (pp. 292-296).

De Wever, B., Schellens, T., Valcke, M., & Van Keer, H. (2006). Content analysis schemes to analyze transcripts of online asynchronous discussion groups: A review. *Computers & Education, 46*(1), 6-28.

Derry, S. J., Pea, R. D., Barron, B., Engle, R. A., Erickson, F., Goldman, R., . . . Sherin, B. L. (2010). Conducting video research in the learning sciences: Guidance on selection, analysis, technology, and ethics. *Journal of the Learning Sciences, 19*(1), 3-53.

Hakulinen, L. (1979). *Suomen kielen rakenne ja kehitys* (4th ed.). Helsinki, Finland: Otava.

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in english*. London: Longman.

Harteis, C. (2018). Machines, change, work: An educational view on the digitalization of work. In C. Harteis (Hrsg.), The impact of digitalization in the workplace – an educational view (S. 1-10). Dordrecht: Springer.

Hämäläinen, R., & De Wever, B. (2013). Vocational education approach: New TEL settings—new prospects for teachers' instructional activities? *International Journal of Computer-Supported Collaborative Learning, 8*(3), 271-291.

Härkönen, R. (2016). *Tilanteen vaikutus 14-vuotiaiden puheen akustisiin ja perkeptuaalisiin piirteisiin (acoustical and perceptual analysis of the situational effect on 14-year-olds' speech)*

Hellermann, J. (2003). The interactive work of prosody in the IRF exchange: Teacher repetition in feedback moves. *Language in Society, 32*(1), 79-104.

Imhof, M., Välikoski, T., Laukkanen, A., & Orlob, K. (2014). Cognition and interpersonal communication: The effect of voice quality on information processing and person perception. *Studies in Communication Sciences, 14*(1), 37-44.

Isohätälä, J., Järvenoja, H., & Järvelä, S. (2017). Socially shared regulation of learning and participation in social interaction in collaborative learning. *International Journal of Educational Research, 81*, 11-24.

Kankare, E., Laukkanen, A., Ilomäki, I., Miettinen, A., & Pylkkänen, T. (2012). Electroglottographic contact quotient in different phonation types using different amplitude threshold levels. *Logopedics Phoniatrics Vocology, 37*(3), 127-132.

Kiuru, N., Poikkeus, A. M., Lerkkanen, M. K., Pakarinen, E., Siekkinen, M., Ahonen, T., & Nurmi, J. E. (2012). Teacher-perceived supportive classroom climate protects against detrimental impact of reading disability risk on peer rejection. *Learning and Instruction, 22*(5), 331-339.

Kumpulainen, K., & Lipponen, L. (2010). Productive interaction as agentic participation in dialogic enquiry. In K. Littleton, & C. Howe (Eds.), *Educational dialogues: Understanding and promoting productive interaction* (pp. 48-63). London: Routledge.

Laukkanen, A., Vilkman, E., Alku, P., & Oksanen, H. (1997). On the perception of emotions in speech: The role of voice quality. *Logopedics Phoniatrics Vocology, 36*(3), 465-475.

Laver, J. (1991). Voice quality and indexical information. In J. Laver (Ed.), *The gift of speech. papers in the analysis of speech and voice* (pp. 147-161). Edinburgh: Edinburgh University Press.

Lehiste, I. (1970). *Suprasegmentals*. Cambridge, Massachusetts: MIT Press.

Lemke, J. L. (1990). *Talking science: Language, learning, and values.* New Jersey, USA: Ablex Publishing Corporation.

Leszcz, M. (2017). How understanding attachment enhances group therapist effectiveness. *International Journal of Group Psychotherapy, 67*(2), 280-287.

Linell, P. (1998). *Approaching dialogue : Talk, interaction and contexts in dialogical perspectives*. Amsterdam ; Philadelphia, PA: J. Benjamins Pub. Co.

Littleton, K., & Mercer, N. (2013). *Interthinking: Putting talk to work*. London: Routledge.

Lukkarila, P., Laukkanen, A., & Palo, P. (2012). Influence of the intentional voice quality on the impression of female speaker. *Logopedics Phoniatrics Vocology, 37*(4), 158-166.

Lyberg-Åhlander, V., Haake, M., Brännström, J., Schötz, S., & Sahlén, B. (2015). Does the speaker's voice quality influence children's performance on a language comprehension test? *International Journal of Speech-Language Pathology, 17*(1), 63-73.

Lyons, J. (1977). *Semantics*. Cambridge, UK: Cambridge University Press.

Mehan, H. (1978). Structuring school structure
. *Harvard Educational Review, 48*(1), 32-64.

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

Mercer, N., & Dawes, L. (2014). The study of talk between teachers and students, from the 1970s until the 2010s. *Oxford Review of Education, 40*(4), 430-445.

Mercer, N., Dawes, L., & Staarman, J. K. (2009). Dialogic teaching in the primary science classroom. *Language and Education, 23*(4), 353-369.

Mercer, N., Wegerif, R., & Dawes, L. (1999). Children's talk and the development of reasoning in the classroom. *British Educational Research Journal, 25*(1), 95-111.

Mortimer, E., & Scott, P. (2003). *Meaning making in secondary science classrooms*
(1st ed.). Berksire, England: Open University Press.

Muhonen, H., Rasku-Puttonen, H., Pakarinen, E., Poikkeus, A., & Lerkkanen, M. (2017). Knowledge-building patterns in educational dialogue. *International Journal of Educational Research, 81*, 25-37.

Nassaji, H., & Wells, G. (2000). What's the use of 'triadic dialogue'?: An investigation of teacher-student interaction. *Applied Linguistics, 21*(3), 376-406.

Niemi, K. (2016). *Moral beings and becomings: Children's moral practices in classroom peer interaction*

O'Connor, C., & Michaels, S. (2007). When is dialogue 'Dialogic'? *Human Development, 50*(5), 275-285.

O'Connor, J. D., & Arnold, G. F. (1973). *Intonation in colloquial english*
(2nd ed.). London: Longman.

Oksanen, K., & Hämäläinen, R. (2010). (2010). Using psychophysiological methods in the research of collaborative learning games. In Anonymous (Eds.) *International Symposium on Collaborative Learning and Argumentation (ICLA 2010),*

Pépiot, E. (2013). Voice, speech and gender: Male-female acoustic differences and cross-language variation in english and french speakers. *XVèmes Rencontres Jeunes Chercheurs De l'ED 268, HAL ID: halshs-00764811*

Postareff, L., Mattsson, M., Lindblom-Ylänne, S., & Hailikari, T. (2017). The complex relationship between emotions, approaches to learning, study success and study progress during the transition to university. *Higher Education, 73*(3), 441-457.

Rogerson, J., & Dodd, B. (2005). Is there an effect of dysphonic teachers' voices on children's processing of spoken language? *Journal of Voice, 19*(1), 47-60.

Routarinne, S. (2003a). Parenteesit ja nouseva sävelkulku keskustelun kielioppiin. *Virittäjä, 107*(3), 398.

Routarinne, S. (2003b). *Tytöt äänessä. parenteesit ja nouseva sävelkulku kertojan vuorovaikutuskeinoina*. Helsinki, Finland: Suomalaisen kirjallisuuden seura.

Salloum, S., & BouJaoude, S. (2017). The use of triadic dialogue in the science classroom: A teacher negotiating conceptual learning with teaching to the test. *Research in Science Education,*

Scherer, K. R. (1972). Judging personality from voice: A cross-cultural approach to an old issue in interpersonal perception. *Journal of Personality, 40*(2), 191-210.

Scherer, K. R., & Giles, H. (1979). *Social markers in speech*. Cambridge, UK: Cambridge University Press.

Sinclair, J. M., & Coulthard, R. M. (1975). *Towards an analysis of discourse: The english used by teachers and pupils*
(1st ed.). London: Oxford University Press.

Vilkman, E., & Manninen, O. (1986). Changes in prosodic features of speech due to environmental factors
. *Speech Communication, 5*(3-4), 331-345.

Vygotsky, L. S. (1987). Thinking and speech
. In R. W. Rieber, & A. S. Carton (Eds.), (1st ed.,). New York: Plenum.

Waaramaa, T. (2014). Perception of emotional nonsense sentences in china, egypt, estonia, finland, russia, sweden, and the USA. *Logopedics, Phoniatrics, Vocology, 40*(3), 129-135.

Waaramaa, T., Alku, P., & Laukkanen, A. (2006). The role of F3 in the vocal expression of emotions. *Logopedics Phoniatrics Vocology, 31*(4), 153-156.

Waaramaa, T., Laukkanen, A., Airas, M., & Alku, P. (2010). Perception of emotional valences and activity levels from vowel segments of continuous speech. *Journal of Voice, 24*(1), 30-38.

Waaramaa, T., Palo, P., & Kankare, E. (2014). Emotions in freely varying and mono-pitched vowels, acoustic and EGG analyses. *Logopedics Phoniatrics Vocology, 40*(4), 156-170.

Waaramaa, T., & Kankare, E. (2013). Acoustic and EGG analyses of emotional utterances. *Logopedics Phoniatrics Vocology, 38*(1), 11-18.

Waaramaa, T., & Leisiö, T. (2013). Perception of emotionally loaded vocal expressions and its connection to responses to music. A cross-cultural investigation: Estonia, finland, sweden, russia, and the USA. *Frontiers in Psychology, 4*, 344.

Warwick, P., Vrikki, M., Vermunt, J. D., Mercer, N., & van Halem, N. (2016). Connecting observations of student and teacher learning: An examination of dialogic processes in lesson study discussions in mathematics. *Zdm, 48*(4), 555-569.

Zald, D. H. (2003). The human amygdala and the emotional evaluation of sensory stimuli (review). *Brain Research Reviews, 41*, 88-123.

Zellner Keller, B. (2004). Prosodic styles and personality styles: Are the two interrelated? In Anonymous (Eds.)*Proceedings of the 2nd International Conference on Speech Prosody – SP2004 International Conference on Speech Prosody – SP2004,* (pp. 383-386).

Appendix

dB: http://science.howstuffworks.com/question124.htm. Dec 12th 2016.

Frequency: https://www.merriamwebster.com/dictionary/frequency#medicalDictionary. Dec 12th 2016. ?

Fundamental frequency (F0 and pitch):
https://www.researchgate.net/post/What_is_Pitch_or_Pitch_Frequency_of_a_speech_signal. Dec 12th 2016.?

Hertz (Hz): https://www.merriam-webster.com/dictionary/hertz#medicalDictionary. Dec 12th 2016.

Loudness: http://hyperphysics.phy-astr.gsu.edu/hbase/Sound/loud.html. Dec 12th 2016.?

Pitch: https://www.merriam-webster.com/dictionary/pitch#medicalDictionary. Dec 12th 2016.

Prosody: https://www.merriam-webster.com/dictionary/prosody. Dec 12th 2016; ?
http://grammar.about.com/od/pq/g/prosodyterm.htm. Dec 14th 2016.

Semi-tone: https://www.merriam-webster.com/dictionary/semi-tone. Dec 12th 2016.