

## Research-based professional development workshops for EFL teachers: Focus on oral test interaction and assessment

Erica Sandlund<sup>1</sup>, Pia Sundqvist<sup>2</sup>, Lina Nyroos<sup>3</sup>

<sup>1</sup>*Department of Language, Literature, and Intercultural Studies, Karlstad University, SE-65188 Sweden, +46 54 7001709, [erica.sandlund@kau.se](mailto:erica.sandlund@kau.se)*

<sup>2</sup>*Department of Language, Literature, and Intercultural Studies, Karlstad University, SE-65188 Sweden, +46 54 7001508, [pia.sundqvist@kau.se](mailto:pia.sundqvist@kau.se)*

<sup>2</sup>*Department of Scandinavian Languages, Uppsala University, SE-75105 Sweden, +46 18 4717873, [lina.nyroos@nordiska.uu.se](mailto:lina.nyroos@nordiska.uu.se)*

### Abstract

In this paper, we address language teachers' continuing professional development (CPD) and engagement with research by demonstrating how research on English as a foreign language (EFL) speaking tests can be applied in professional training workshops on oral test interaction and assessment. Data were drawn from an ongoing research project targeting 9<sup>th</sup>-grade national English speaking tests in Sweden. Authentic test recordings analysed in a previous study (Sandlund and Sundqvist 2013) formed the basis for workshops designed to initiate discussions on testing and assessing EFL oral proficiency. The two-fold aim of the study is to implement a research-based workshop model for increasing teachers' awareness of interaction in speaking tests, and to analyse teachers' self-reported insights from workshop participation with regard to conducting speaking tests and assessing oral proficiency. Workshop experiences were followed up with questionnaires. Findings reveal some noteworthy pedagogical implications of collaborative work with test recordings and assessment.

Keywords: EFL; speaking test; CPD; assessment; test interaction

### 1. Introduction

Professional development and life-long learning for teachers at all levels of educational systems are concerns for politicians and educational institutions as well as for teachers and researchers. Central issues in research on teachers' in-service training and professional development include how educational reform and societal change increase the needs for continuous training for teachers (Day and Sachs 2004a), the effectiveness of various models for teacher development (Bolam and McMahon 2004), and the ways in which new educational or subject-matter research may be successfully incorporated into in-service training and practice (Erlam 2008). In the area of language teaching and learning, it has been

argued that the gap between new research on the one hand, and teachers' daily practice on the other, is perhaps particularly wide (e.g. Ellis 1997, Nassaji 2012). Questions for debate in relation to the research/practice divide include: *i*) what are some productive ways of making second or foreign language research accessible and relevant to teachers, teacher education, and in-service training, and *ii*) what formats of professional development are demonstrably successful in engaging language teachers, in and with current research, in ways that have positive effects on language classroom practices? Language teachers are expected to teach their students a number of different skills, including the 'communicative language competence' activated in the performance of oral and written reception, production, interaction or mediation as described in the *Common European Framework of Reference for Languages* (Council of Europe 2001). In light of such expectations, it seems inevitable that formats for teachers' continuing professional development should reflect this multi-layered practice. While some theories and models for training and research engagement may be suitable for, for example, improving language teachers' assessment literacy for second/foreign language writing, they may be unsuitable for others, such as learners' ability to interact in a second/foreign language.

In this paper, we describe and analyse the process and outcome of a research-based training workshop for English teachers in Sweden. We build on a recently published empirical study on interaction in EFL oral proficiency tests (Sandlund and Sundqvist 2013) and present the next step in an interventionist research effort: the route from analysis and published research to its practical application in professional development events for language teachers. Through showing how data and findings from a research study are used as a basis in professional development workshops, we aim to address both questions of language teachers' professional development mentioned above, namely language teachers' engagement with research and methods for making such research accessible in training sessions. The aim of the current study is, thus, to test an interventionist approach from empirical research to language teachers' professional development, using published research as a basis in researcher-mediated workshops. Research questions for the present study are:

- *RQ1*) Can data and findings from an empirical study on social interaction in second/foreign language oral proficiency tests be successfully used in professional development sessions on the testing and assessment of oral proficiency?
- *RQ2*) What are some of the insights and expressed learning outcomes for language teachers participating in such a research-based workshop?

Using data from five workshops conducted as part of a large research project on high-stakes EFL tests [Testing Talk, reg. no 2012-4129], including post-workshop questionnaires, we discuss how the use of authentic research data, research findings from a study on social interaction in EFL tests, and collegial discussions during workshops can constitute a specialized format for teachers' in-service training. The workshop format aims to provide an opportunity for teachers to engage directly with current research, and also aims to create a space for reflection upon oral proficiency testing and assessment. We begin with a brief

overview of research on teachers' continuing professional development and research engagement, after which we take readers through one specific workshop exercise used on five different workshop occasions. Finally, we examine questionnaire data from these five workshop occasions, and discuss some pedagogical implications of the workshop setup.

## 2. Research on language teachers' professional development

Training for in-service teachers is often discussed under the umbrella term continuing professional development (CPD). Very generally, CPD refers to 'all the activities in which teachers engage during the course of a career which are designed to enhance their work' (Day and Sachs 2004b, 3), even though it has been argued that professional development for teachers is 'more inclusive of personal and moral dimensions' as compared to career-oriented CPD of other professional groups (Mann 2005, 104). Central goals of teachers' CPD mentioned in the literature include 1) aligning teacher practices with educational policy, 2) improving learner performance by improving teacher performance, and 3) enhancing the status and profile of the teacher profession (Day and Sachs 2004a, 22).

In a paper on factors influencing language teachers' engagement in and with research, Borg (2010, 412) discusses the common view that published research 'should directly inform practice', and instead proposes that second/foreign language acquisition research should be viewed as a 'source of enhanced understanding of their work, not as a direct solution to their problems' (419). A number of epistemological and practical conditions deemed necessary for teachers' productive engagement with and in research are listed, for example that 'collaborations among teachers, and among academics and teachers create productive mutually-beneficial social spaces for knowledge creation' (418). In addition, Borg discusses the translation process from published research to something that is accessible to teachers, arguing that research in the form of 'detailed case studies of aspects of classroom life, may represent language learning and teaching phenomena to teachers in ways which they can relate more immediately to their own experience' (416). For teachers to be not only engaged in research projects as informants, but to also be 'critical consumers of educational research, using it to inform their instructional decisions' (410), published research needs to be made available and accessible. Moreover, teachers need to be motivated to read recent research, and they also need the time and skills to read it. Recently, Borg (2015) has developed his ideas regarding language teachers' professional development further and, among other things, he suggests seven characteristics for what constitutes effective professional learning: (i) relevant to teacher and student needs, (ii) involvement of teachers in decisions about content and process, (iii) collaboration between teachers and sharing of expertise, (iv) a collective enterprise with support from schools/educational systems, (v) "exploration and reflection are emphasized over methodological prescriptivism", (vi) availability of internal and/or external support, and (vii) teachers' own inquiry is viewed as a learning process. Most, if not all, of these characteristics are applicable to the workshop CPD format presented below (Borg 2015, 6).

In a conceptual discussion concerning teachers' professional knowledge, Cochran-Smith and Lytle (1999) discuss three types of knowledge as regards development for teachers: knowledge *for*, *of*, and *in* practice. They argue that '[d]ifferent conceptions of teacher learning—although not always made explicit—lead to very different ideas about how to improve teacher education and professional development' (Cochran-Smith and Lytle 1999, 249). Further, Lieberman (1996) identifies three contexts where professional development can take place: *direct learning* (conferences, consultations, workshops etc.), *learning in school* (peer-driven learning, mentoring, team planning and assessment, research activities etc.), and *learning out of school* (professional networks, partnerships between schools/universities etc.). It has been argued that teachers will need to 'engage with all of these if they are to keep pace with and respond to changes in society, the demands of the result-driven standards agendas of governments and at the same time retain their energy, enthusiasm and commitment to high-quality teaching' (Day and Sachs 2004b, 13).

In reality, however, many (language) teachers report increased administrative duties, work overload (cf. Smith and Ingersoll 2004), and a lack of professional development in their own subject areas due to costs and to priority being given to general topics related to educational reform (National Union of Teachers in Sweden 2006). Smith and Ingersoll (2004) show that novice teachers who receive the opportunity to participate in collective activities, such as planning and collaborating with their colleagues, are more likely to stay in the profession than beginning teachers who are not engaged in similar in-service activities. In Sweden, results from several reports in the 2000s (see, e.g., National Union of Teachers in Sweden 2006) indicated that there was an imminent need for in-service training among teachers; in particular it was considered necessary to offer CPD courses targeting assessment. Therefore, the government introduced a national CPD program (Boost for Teachers) over a period of three years (2007–09), in which 11,000 teachers participated. The program was evaluated and participants were very positive (State Council 2010). Certain teacher categories were given priority to participate in the Boost program; language teachers (including English teachers) were however not a select teacher category (State Council 2010). Thus, in Sweden, there is a great need for CPD courses that target or involve language assessment in a school subject such as English. This includes all levels of language assessment, including oral proficiency. However, what kinds of preparation do English teachers need in order to provide equitable assessments of learners' oral proficiency in a second/foreign language? As research on the testing and assessment of second/foreign language oral proficiency has shown (e.g., Kasper and Ross 2007, Galaczi 2008, Nakatsuhara 2008, Lazaraton and Davis 2008, Sandlund and Sundqvist 2011), attention to social interaction among test-takers (and examiners) during tests is significant for test constructors, examiners, and teachers, and such research has also pointed to the interplay between interactional patterns and individual assessments assigned to test-takers.

The present research project had an explicit aim of identifying and testing a CPD format that would allow insight into the complex weave of social interaction in an oral language testing context and where issues of oral proficiency assessment can be discussed, but which differed from traditional rater training. Instead, the sought-after CPD format would begin *prior* to assessment, in views and preconceptions about test interaction, task management, and

examiner conduct. Although assessment is often discussed in workshops, the goal is not for participating teachers to arrive at the same assessment, but rather, to problematize and discuss issues that may play a role in assessing test-takers individually on a product that is inherently jointly constructed. Our review of the CPD literature for second language teachers revealed a lack of empirical studies as well as training efforts where teachers could participate in researcher-mediated reflections upon authentic test conduct in recordings. By way of comparison, courses on, for example, assessment theory and practice (offered as part of Boost for Teachers in Sweden) generally leave out interactional perspectives on oral testing and assessment.

Looking beyond CPD for teachers and toward other professions, a research-based workshop method for communication skills training for professionals was developed by conversation analyst Stokoe (2014, 2011): the Conversation Analytic Role-Play Method (CARM), which builds on Stokoe's work on calls to community mediation services. In CARM workshops, recordings and detailed transcription of turns in authentic mediation calls are synchronized when played in a slide show presentation, so that 'workshop participants "live through" the call as it happens' (Stokoe 2011, 126). The recording can be played line-by-line with opportunities for discussion and role-playing of next possible responses for each turn. When role-playing, participants can draw on and reflect upon their professional experience and practice before the authentic continuation is resumed. In a recent paper, Stokoe (2014, 255) describes how CARM is designed as 'a unique framework for discussing and evaluating, in slow motion, actual talk as people do their jobs' and 'provides an evidence base for making decisions about effective practice and communication policy in organizations' (see also Jenkins and Reuber 2014, for an application of the method in the healthcare sector).

The workshop format developed in the present study was inspired by Stokoe's (2011) use of authentic recordings and application of interaction analyses in workshops with mediators. However, given that the role of the language teacher in oral proficiency tests differs from that of professionals in other types of talk-based institutional interactions (that is, in the national tests in question, the teacher is instructed to remain relatively passive in the test-takers' conversations, although actual practice varies), the explicit communication skills training approach and role-playing aspects of CARM were deemed less suitable for our target audience. Instead, we developed training sessions designed to be valuable specifically for teachers administering and assessing second/foreign language oral proficiency tests. The workshops target teachers' understanding of test-taker conduct and the assessment of particular aspects of their interactional conduct. Our workshops are built on the idea of bridging the well-documented gap between second/foreign language research and language pedagogy, and attends to the issue of making second/foreign language research more *accessible* to in-service teachers (cf. Ellis 1997, Erlam 2008, Borg 2010, 2015). Through making available and discussing recordings from authentic high-stakes oral proficiency tests, a particular phenomenon such as oral proficiency testing, which language teachers can relate to in their work, is made available for collegial and researcher/teacher discussions. The model also aims to incorporate the types of knowledge described by Cochran-Smith and Lytle (1999) in that it involves theories of language learning and testing, while also specifically inviting teachers to reflect upon their own practice and evaluate the practices of others. In turn, the

training sessions aim at yielding insights that inform each participating teacher's inquiry into their own testing and assessment practices. The model proposed here also constitutes a blend of collegial learning and direct learning through dialogue between second and foreign language scholars and in-service teachers (cf. Lieberman 1996), and aims to initiate collegial discussions and collaboration on oral language testing and assessment. In addition, workshops serve as a venue for research communication to language teachers.

### 3. Data and methods

Data for the present paper consist of 1) research-based workshop materials from one of the workshop exercises ('History'), 2) ethnographic field notes from discussions at five workshops, and 3) post-workshop questionnaires ( $N = 54$ ) administered at the end of each workshop. Moreover, the workshop materials consist of recordings, transcripts, and analyses from an empirical study published by two of the authors in this journal (Sandlund and Sundqvist 2013). The data for that particular study consists of recordings from the speaking part of the national test of English as a foreign language for ninth graders in Sweden, a high-stakes, summative test in which the test-taker's own teacher functions as the test administrator, instructor, and examiner. Internationally, it is uncommon to have high-stakes speaking tests in which the test-takers' own teacher is the examiner, but some countries adopt such a procedure, for example, New Zealand (East 2014), Norway (Hasselgren 2000), and Sweden. The Swedish case constitutes a particularly interesting case in point since the Swedish testing system combines a high level of faith in teachers (scoring their own students) with the characteristics of a typical accountability system (Lundahl and Tveit 2014) (for additional details about the national test of English in Sweden, see Erickson 2012, Sandlund and Sundqvist 2011).

The workshops were conducted as part of an interventionist segment of the Testing Talk project, funded by Committee for Educational Sciences, Swedish Research Council. The Committee encourages projects with 'a strong connection to the education of teachers' and which contribute to the 'scientific basis of this education, and to the future development of schools' (<http://www.vr.se>).

#### 3.1 Research underlying workshops

As mentioned, the basis for the training workshops is our research, beginning in 2010. Paired oral proficiency tests were analysed using Conversation Analysis (CA), an established approach to the study of human social action in talk and embodied conduct (Sidnell and Stivers 2013). Originally a sociological approach, CA uncovers the systematics of talk-in-interaction through analysis of recordings and the production of detailed transcripts of talk and action. The workshop training aligns with what Antaki (2011, 8) labels *interventionist CA*: conversation analytic research that explicitly aims to contribute to finding solutions to pre-existing institutional problems.

As for ideas to the workshop materials, one of our studies had revealed some inconsistencies with regard to how individual test-takers had been assessed when comparing interactional patterns with assessment data (see Sandlund and Sundqvist 2011). For example, in comparing the analyses of interaction between test-takers to subsequent assessments, Sandlund and Sundqvist showed that certain task management strategies were associated with higher grades than others, despite the fact that such strategies were, in fact, productive for the test-takers in managing upcoming problems *in situ*. These findings raised a question of which aspects of test-takers' interactional conduct that come into play in the final assessment. Recently, Kasper (2013, 260) claimed that this study is 'the only effort [so far] to investigate how test takers orient to the task instruction in peer oral assessment, and how such task orientations are related to rating outcomes' (see also Okada and Greer 2013). Furthermore, and particularly important for the present paper, is another study based on the same national test of English, which revealed how teachers' and test-takers' understanding of test tasks diverged, which had consequences for the outcome of the testing interaction (see Sandlund and Sundqvist 2013). The workshop exercise selected for the present paper is taken from the 2013 article, and shows how two test-takers and one teacher are treating the pre-set topic for discussion rather differently. We will return to some specifics of those findings and how they facilitated teachers' reflections when we account for the workshop exercise in section 4 below.

### ***3.2 Participants and post-workshop questionnaires***

Participants in workshops are all English teachers in Sweden. Five workshops with the same training materials were conducted, and three of these were conducted at schools participating in the Testing Talk project. These teachers had all consented to participating in a research project where additional national test data was to be collected, and they had been informed beforehand that they would be offered to participate in two different workshops during the course of three years. These teachers were employed at four secondary schools in southwestern Sweden; two were located in a medium-sized city and two in rural areas. Teachers at the two latter schools were joined in the same workshop. All teachers were female and all had some or extensive experience of administering the national test of English in the ninth grade.

In addition, two workshops were conducted as open events at the university, as part of an annual conference event organized by researchers at the university and designed for teachers of languages and literature in the region. Invitations to the workshops were sent out as part of the general conference invitation via email lists to schools in the region. Participation in the workshops was optional as workshops were offered before lunch, and lectures after lunch. Participants in the two university-based workshops were informed that the workshop would focus on the national test of English and on interaction in such tests, and signed up to participate. All participants reported to be teachers of English.

In order to access any perceived learning outcomes from participating in a workshop, teachers

were asked to fill out a questionnaire at the end of each workshop. It included four open-ended questions, and for the purpose of this paper, responses to the first question ('Are there insights from today's workshop that you think will inform your own teaching and/or testing practices?') were examined using content analysis (Cohen, Manion, and Morrison 2007). The three school-based workshops had between five and nine teachers, and the open workshops drew around 40 teachers, but not all of them turned in the questionnaire. In total, 54 questionnaires were collected from the five workshops.

### *3.3 Developing the speaking test workshop*

In preparing workshops, we used peer-reviewed research and the accompanying data: audio recordings of authentic high-stakes speaking tests (from a corpus of 199 EFL tests, see Sundqvist 2009). All recordings used for the workshops are from a previous national test of English.<sup>1</sup> For the present paper, we focus on one workshop exercise ('History'), used in all five workshops and based on an analysis from one of our publications (see Sandlund and Sundqvist 2013 for a detailed analysis). This particular sequence was selected over others for several reasons: for one, this particular sequence makes possible a discussion of the teacher's role in paired speaking test, which was deemed a relevant topic to the teachers. Furthermore, an interesting feature of the sequence was the teacher's and test-takers' diverging interpretations of the test topic, as displayed in their treatment of the task and prior turns. Finally, the sequence is long enough to provide several opportunities for addressing what 'oral proficiency' is in terms of interactional strategies, language use, and topic management.

For workshops, a slide show presentation was prepared where transcripts of selected sequences are displayed on a screen and where the sound files have been synchronized with the appearance of each line in the transcript (cf. Stokoe 2011). While we always begin by allowing the groups of teachers to discuss freely a recording segment, we also prepare questions for each sequence, so that teachers' initial responses to what they hear can be contextualized and gradually more focused on particular issues. Letting workshop participants hear and discuss the interaction as it develops on a turn by turn basis is also consistent with the way our research was conducted: by focusing on the details of each turn and its immediate responses, CA as a method is designed to uncover interactants' own interpretations of the ongoing talk, displayed in their treatment of preceding actions in designing next turns. Such an emic (members') perspective on talk (see, e.g., Markee and Kasper 2004) is directly transmitted from our analysis of data to the workshop setup.

As part of a mission from the government, the Swedish Schools Inspectorate (2012, 2013) has published several reports where the national tests are subject to criticism; one reason being that the scores differed (sometimes extensively) when tests were re-assessed by others than the students' own teachers. A heated public debate on national tests and their assessment

---

<sup>1</sup> Test tasks from the 2007 test are freely available at [http://nafs.gu.se/prov\\_engelska/exempel\\_provuppgifter/engelska\\_ak9\\_exempelupp](http://nafs.gu.se/prov_engelska/exempel_provuppgifter/engelska_ak9_exempelupp)

followed, which spurred us to also target equity in assessment in our workshops (see also Gustafsson and Erickson 2013). Therefore, assessment of oral proficiency was also discussed in relation to each exercise. Assessments of a particular test-taker dyad/group from their original teacher and three expert raters were shared and discussed at the end of each exercise. In this way, teachers' perceptions of what qualities of test interaction to focus on from an assessment perspective are unveiled, which in turn may contribute to improving the validity of the test (cf. Messick 1989). However, because of space limitations, the assessment modules will not be detailed in this paper. Also, since authentic assessments are based on full-length tests, and workshops centre on selected segments and themes, assessment discussions are not to be considered entirely representative of teachers' practice, but the recordings open up for rich opportunities for problematizing what oral proficiency 'is' and 'should be', and which aspects of interactional conduct may come into play.

Depending on the size of the workshop group, participants are divided into small groups with 3–4 participants in each. After having played each recording segment, the groups are allowed discussion time (10 minutes or more) before a plenum discussion with the researchers takes place. The workshops are researcher-led, but all workshop groups always offer their observations and thoughts prior to the researcher-led discussions. After the groups have shared their thoughts and we have discussed different opinions, some research-based observations are shared by the research team. Inevitably, researcher-led workshops such as the ones discussed in this paper have their limitations. One argument against this setup is that priority is given to the researchers' perspective in the concluding discussion of each recording and that teachers' views and experiences are contrasted with the published findings. Also, since the workshop training is based on the researchers' own research, a concern about a one-sided perspective on testing interaction may be raised. A few arguments for the setup may nevertheless be offered in return. Firstly, our strong belief in the authenticity of training materials for effective CPD requires access to authentic test data. Given that the workshops centres on the national test with real test-takers in a high-stakes testing context, a collecting of recordings for training purposes would nevertheless have to obtain consent from participants and ethical approval. As such, research projects with an explicit interventionist aim are a suitable vantage point for designing such materials. Secondly, given that most research may require some 'translation' or specialized methodological knowledge in order to be accessible and applicable to practitioners, researcher presence is considered an advantage rather than a disadvantage. Since the workshop leaders themselves have worked extensively on testing interaction data and conducted studies that form the basis of the workshop, they are in a unique position to communicate authentic research findings, and to present them in relation to other scholars' work on the topic. It would, in essence, be very difficult for a non-CA researcher to lead a workshop based on other CA scholars' findings on test interaction, and most of the current efforts to use CA research interventionistically involve the original researcher (cf. Antaki, 2011). In order to give workshop participants a broad overview of current research on oral language testing and social interaction and not only the research team's own perspective, each workshop begins with a short lecture and research overview. Teachers are also supplied with references to all works cited.

## 4. The workshop study

### 4.1 The 'History' exercise

This section gives a hands-on demonstration of the 'History' exercise along with ethnographic field notes of teachers' discussions. The particular recording used was collected from the national speaking test in 2007 ([http://www.nafs.gu.se/digitalAssets/1356/1356787\\_theworldaroundus.pdf](http://www.nafs.gu.se/digitalAssets/1356/1356787_theworldaroundus.pdf)). In this test, test-takers are instructed to discuss statements or questions such as 'Money makes people happy' or 'What can we learn from history?' from topic cards.

In a workshop, we go through three or four exercises with different thematic foci. As a workshop begins, participants are divided into small groups. The 'History' exercise is based on a test sequence presented across five slides and divided into ten audio clips, where each clip is followed by discussion time. The sequence is 2 minutes and 23 seconds long (test duration: 12 minutes, 23 seconds) and interaction participants are an EFL teacher (TEA in transcript) and two test-takers (PER and LIV). We have observed discussion times up to 90 minutes on this segment alone as it encompasses multiple aspects of test interaction and assessment. Due to space limitations, however, we address only a selection of the issues explored in the workshops. All details of the CA transcription (see Appendix) are preserved, thereby maintaining the authenticity of the research data. Scissor and star icons indicate clips/stops and points of particular relevance for discussion of assessment. Figures 1–10 illustrate the data transcripts displayed on the screen. For a full analysis of the sequence, see Sandlund & Sundqvist (2013) in this journal.

The extract shows the initiation of one topic card, formulated as a question for the test-takers to discuss. The sequence topic is 'What can we learn from history?', and in workshops, we follow the interaction up to when another card is drawn, representing a complete stretch of talk for a topical task. We begin with the first audio clip and transcript (Figure 1).

- 36 LIV *den?*  
*this one?*
- 37 (6.6) ((looking through the stack of topic cards))
- 38 LIV (pt) '↑*happ* u::h (.) what can we <le:arn> from  
*okay*
- 39 history?
- 40 (2.1)
- 41 LIV ((sniveling sound))
- 42 LIV °(hmm)°
- 43 LIV history y'know the:: u:h (1.5) *ämnet* in the school or (.) .hh ★1  
*the subject*
- 44 (1.2)
- 45 LIV hist[o r y ]



Figure 1. Lines 36–45 .

We play the clip from line 36 through 45 a few times and ask the groups to freely discuss what they hear. Their initial responses usually deal with everything from how the test topic is formulated to the two instances of codeswitching: ‘*happ* (condensed form of Swedish *jaha*, here *okay*) and *ämnet* (*the school subject*). They also discuss how slight deviations from the “speak English only” instructions may influence assessment (★1). Having discussed the teachers’ own observations, we ask them to identify the problem that Liv appears to run into, and to discuss possible reasons for Liv’s question (lines 43–44). After 5–10 minutes, we ask the groups to share their thoughts. Teachers immediately identify Liv’s problem as related to the task (whether history refers to the school subject or history in general) and most participants express surprise at the complexity of this question. We then ask teachers to speculate about what will happen and what they think they possibly would have done next. Here, we usually receive a great variety of responses – while some argue they would give Per a chance to help Liv out before stepping in, others would try to spur Liv to see world history beyond the school subject, or alternatively, allow Liv to interpret the question in any way she wants. Yet others find it very difficult to speculate at all.

We then play the teacher’s response (line 46) and continue down to line 54 (see Figure 2).

46 TEA <[no his]tory in the< world. (.) [ I ] ( I mean ),  
 47 LIV [ye-] >in the world yeh .hh <  
 48 (1.9)  
 49 LIV thatu:h the world is going for↑ward (.)  
 50 for example for- four hundred years ago we: had .hhh (1.1)  
 51 when it was war we:::u hhTHUH HHH (0.9) we didn't (.)  
 52 shout out guns we:: (.) fight withu:h, ★2  
 53 PER sword  
 54 LIV °yess::°



Figure 2. Lines 46–54.

Participants discuss and evaluate the teacher's initial *no*, which casts Liv's first interpretation as erroneous, and continue discussing Liv's actual topic card response and examples of what we learn from history. At this point, we ask participants to evaluate Liv's response from their own professional vantage point. We also take a close look at the details of her turn production, discussing whether content production (that is, Liv's choice of example, in this case warfare) in relation to assessment. They talk about what constitutes fluency and reflect upon the errors of Liv's turn, such as *shout out with guns* and the preposition use in *for four hundred years ago* (a direct transfer from Swedish syntax) (★2). In the end, most participants agree that Liv adheres to the instructions: she is producing coherent English and relating her response to the topic card. We also discuss Per's minimal contribution *sword* (line 53), and although most teachers note that Per has been passive so far, they also agree that his contribution is properly fitted into the sequential context (helping his co-interactant in a word-search).

We continue to lines 55–62 (Figure 3), discussing how to deal with silences in speaking tests (lines 60 and 62), and whether pauses of a certain length are automatically problematic from an oral proficiency perspective (★3), or if they actually mirror patterns in naturally occurring conversations, which this test format actually aspires to resemble. We also address Liv's second response about cars in relation to topical content and assessment (★4):

55 LIV (.) and now we- (0.2) is not good to have a wa: r but .hh  
 56 now we:: (.) for example shoot with guns ? ★3  
 57 PER HHH[(HH)  
 58 LIV [I (k(hh)no(hh)w?)HHHUH (.) \$and (.) for four hundred years ★4  
 59 we\$ didn't havu:h (.) cars? and now we have cars? .hhh  
 60 (1.9)  
 61 LIV hhmm?  
 62 (5.1)



Figure 3. Lines 55–62.

We then play only the teacher's turn (Figure 4):

63 TEA but what can we ↑learn (.) from history



Figure 4. Line 63.

Participants usually pick up on the initial *but* and the teacher's intonation rise on *learn* and hear the turn as a dismissal of Liv's contributions to the topic, as if Liv has not responded adequately to the topic card question, which is in line with our analysis. We spend quite some time on this turn and try to come up with reasons why the teacher appears to be steering the test-takers away from Liv's focus on technological advancements and also discuss how a curricular agenda may come into play. Workshop participants usually deny that morally or intellectually dubious topical content would make them intervene in this way, but we challenge these perceptions and discuss how topical content may implicitly affect assessment. We also emphasize that the EFL teacher in the recordings is highly experienced and based on the large number of tests recorded, we have no evidence that her conduct is in any way deviating from other teachers in the dataset. In relation to this, we talk with workshop participants about how social interaction relies on tacit and taken-for-granted knowledge of how we *do* conversation; although people believe that they can account for what they actually did in a particular conversation, such accounts are always stripped of their contextual anchoring. In many of the post-workshop questionnaires, teachers report that it was an

enlightening experience to see how seemingly small details of turn production may have major interactional consequences.

Next we address possible response options after the teacher's turn. We ask participants to consider the ways in which the teacher's contribution restricts/promotes particular next actions, and what we can reasonably expect that the test-takers will do. Participants reflect upon the potential problem of stepping in to promote the peer conversation with new angles, which may have the opposite effect on test-takers' independent speech production. The recording is then resumed up to 77 (Figure 5):

64 **LIV** **thatu:h,**  
 65 **(1.2)**  
 66 **LIV** **[yea-]**  
 67 **PER** **<[how ] they lived>?=  
 68 **LIV** **=yeah how they lived. (.) and that we: are going forward** **★5**  
 69 **TEA** **°mmm?° (.) can we e- de- (m) (.) e:hm (.) (pt) (0.3) .hhh do you**  
 70 **get(hhh) (.) any HELp**  
 71 **from knowing that? (0.5) how they lived**  
 72 **LIV** **nnn- (.) no? u::h n:no?**  
 73 **TEA** **°hm?°**  
 74 **LIV** **>Ah no<? hhUH HUH HUH (HHH) [.HHHH**  
 75 **TEA** **[so you think it'sss (0.3)**  
 76 **a progress in wa:r and so on,**  
 77 **(2.1)****

---



Figure 5. Lines 64–77.

We examine Per's response (line 67) and Liv's insistence that the world is *going forward* (line 68) (★5). We also discuss the teacher's insistence on reformulating the 'learning' aspect of human history (69-71), which seems designed to keep Liv's talk close to the topic card question. The teacher's summary of the gist of Liv's contributions (lines 75–76) generated a lively discussion in all five workshops, where participants initially side with the test-taker, arguing that the teacher should give up her attempts to steer the discussion, only later to conclude that they have probably done the same in halting test conversations. Further, assessment is discussed in relation to the diverging interpretations of the topic on the card,

and we encourage teachers to reflect upon their own assessment practice). We then play lines 78-79 (Figure 6) and discuss Liv's response to the teacher's summary and how Liv is essentially rejecting the teacher's analysis of her prior contributions in her hesitant 'nno:':

78 LIV °nno: (.) ( de )° (2.9) ((sniveling sound))  
79 (3.0)

----- ✂

Figure 6. Lines 78–79.

We move on to playing lines 80–87 (Figure 7), followed by 88-93 (Figure 8), and highlight the teacher's question to Per, and the apparent problems in producing his turns (★6, 7):

80 TEA what about Per what do you say.  
81 PER u::hm.  
82 (2.9) ★6  
83 PER .hhhhhh [(HHHHH)  
84 TEA [is it good to know about your history.  
85 (1.8)  
86 PER >yeah<  
87 (0.9)

----- ✂

Figure 7. Lines 80–87.

88 TEA wh(hh)y.  
 89 (0.6)  
 90 PER uh it's (hhhh)  
 91 (6.4) ★7  
 92 PER mmmmmm, .hhh  
 93 (0.4)

----- ✂

Figure 8. Lines 88–93.

We examine the teacher's *why* (line 88) and address how different question formats by teachers allow for different possible responses. Here, we rely upon and give examples from CA research on questions, and contrast yes/no-questions with *wh*-questions. Participants note that a yes/no-question does not promote more talk from Per: structurally, he is only invited to agree or disagree, which he also does. At this point, many teachers find the interaction painstaking. They frequently reflect upon how a peer conversation has turned into something more like a classroom quiz. We continue with the slide illustrated in Figure 9, with the teacher's additional questions (lines 94–99) and participants note that the teacher still insists on reformulating the issue of learning from human history.

94 TEA .hh do we le:arn? from the mistakes we made? .hh  
 95 LIV yes:  
 96 TEA we DO?=  
 97 LIV =hmm  
 98 TEA ( k ) (.) can you tell- (.) give 'xamples  
 99 (3.9)  
 100 LIV well I'm not so good i(HH)n (HH)IS(HH)TOR(HHH)Y (HH)BUT(hh) ★8  
 101 [(HH)HEH (.) HHEH  
 102 TEA [do you have any examples?  
 103 LIV (hh)huh (.) ↑hmmmm↓  
 104 PER >no< (hh)? ★9

----- ✂

Figure 9. Lines 94–104.

For lines 100–104 (Figure 10), we examine how Liv maintains her original interpretation of history as a school subject, and we return to the potential problem of test validity (★8). The teacher’s final attempt to get Per to provide examples fails as he responds with a rapidly produced *no* (★9).

We play the remaining two lines:

105 TEA no? (0.7) .HH ohkay let’s take your then  
 106 ((new topic card introduced))

----- ✂

Figure 10. Lines 105–106.

After going over the topic shift, we end the discussion of this sequence by sharing assessment data for the test-takers, as graded by their teacher and three independent raters (not discussed here). We also problematize the issue of assessing test-takers at an individual level based on conversation that is a joint accomplishment, and when teachers have shared their views and discussed them, we bring the discussion to a close.

## 4.2 *Post-workshop questionnaires*

The ‘History’ sequence was included in all workshops and, as mentioned, in this paper we examine teacher responses to the following item in the questionnaire: ‘What aspects of today’s workshop did you find useful for your own teaching and testing? Please give examples’. Although any self-report data may be flawed in the sense that there is no guarantee that actual changes in professional practice will take place, we argue that the responses indeed reveal new understandings of test interaction and assessment resulting from the training. Some major themes brought up in the comments are discussed below, in relation to the research questions of the study; that is, whether data from interaction research may be productively used in language teacher CPDs, and which insights teachers self-report as learning outcomes of the workshop. Having performed content analysis of 54 item responses, three central learning outcomes emerged from the questionnaire data, described and exemplified in sections 4.2.1–3 (translated from Swedish to English).

### 4.2.1 *Learning outcome 1: The teacher’s role in test preparations and instructions*

The ‘History’ sequence is revealing in that the teacher and test-takers display diverging

understandings of the task and topic, which raises questions as to when to encourage test-takers to abandon a particular test topic. Teachers report insights into the central role of the teacher as a test administrator with regard to providing explicit instructions to the test-takers:

- (1) I feel it's important to prepare the students so they know the test format, to tell them it's okay to take a new topic card.
- (2) It was very interesting to see these interactions in a new light. I myself will definitely be much more explicit in my instructions to students about the purpose of each test occasion.

Similarly, some questionnaire responses expressed an increased awareness of how students may interpret the task, and whether particular interpretations of a topic should matter at all, as long as test-takers produce an intelligible interaction.

#### *4.2.2 Learning outcome 2: Participating in social interaction*

Another theme frequently brought up relates to teachers' insights into the social aspects of participating in speaking tests. In the following responses, teachers report specifically on lessons learned as regards how interaction in this specific setting works:

- (3) It is really important that I as the teacher reflect upon when and why I intervene in test talk so that my interventions bring the conversation forward and not hinder student performance.
- (4) I learned a lot about how recordings can help us understand students' linguistic actions, and about my own role in the conversations – sensitivity and how I formulate questions.
- (5) What happens in a conversation, what silences do, and so forth.

As evidenced above, many teachers reflect upon the workshop data and discussions in relation to their own practice, and reveal insights into ways in which they may be more aware of their own interventions in peer talk in the future.

#### *4.2.3 Learning outcome 3: Collaborative assessment*

In terms of assessment, the workshops allow collegial exchanges of perspectives and experiences of assessing oral proficiency (cf. Borg 2015), spurred by the recordings. Questionnaire accounts bear witness that teachers appreciated the benefits of discussing assessment; workshop discussions brought a realization that many factors not specifically related to oral proficiency may come into play, and that defining what constitutes oral proficiency is not a clear-cut task:

- (6) All teachers assessing speaking should do this. Many things affect how students are assessed, which we need to be aware of. What exactly are we assessing?
- (7) Great workshop format, this could also be used at school conferences for EFL teachers.
- (8) We need to 'up' the discussion of what we assess in speaking tests!
- (9) That oral tests have to be assessed not just by the teacher of the class. That we need to do collaborative assessment by listening together and learning from each other.
- (10) That topic deviations should not be judged so harshly, it is oral proficiency that is central. I also see the benefits of recording oral tests.
- (11) Which aspects of test talk that come into play in assessments – very important for my classroom teaching.
- (12) The feeling that maybe I am too tough when I grade, thus, very important with collaborative discussions for equity in assessments!

As can be seen in quotes (6) – (12), teachers are able to situate the workshop insights in the realm of their own assessment work, and also question what exactly is being assessed. In sum, post-workshop questionnaires indicate that teachers have obtained new insights through working together with the test data.

## 5 Discussion and conclusion

As demonstrated, collaborative work using an extended sequence from a speaking test offers rich opportunities for addressing a multitude of issues of relevance for oral proficiency testing and assessment – such as test preparations, task and topic understandings, silences, and teacher prompting – and their effects on test interaction. The research-based preparation of workshop material allows for discussions that begin with spontaneous reactions, followed by targeted questions from the researchers. As workshop leaders, the researchers can pick up on participants' comments/reactions and lead the discussion toward other possible interpretations, which, in our experience, spurs self-reflection. With regard to research question 1, we deem that the use of authentic data and the particular empirical study was successful; teachers had no problems engaging with the research material or with the preservation of the CA transcription symbols. A tentative interpretation would be that teachers felt a high degree of familiarity with the context and were quickly able to relate to their own practice and experience. In addition, the workshops initiated extended discussions on EFL oral proficiency per se; what exactly does it mean to be a proficient English speaker as a ninth grader in Sweden? Which other qualities, such as morally appropriate opinions, topical familiarity, or a generally slow speaking style, may affect assessments? By allowing teachers to work closely with a piece of empirical data, the workshops, aside from being a forum for sharing research findings, became an opportunity for reflecting upon practice (cf. Borg 2010, 2015). Based on analysis of teacher responses, it is possible to conclude that the new insights are rooted in the close inspection of turns and conduct in testing talk, together with collaborative discussions on assessment. Thus, with regard to research question 2, this study

uncovered three main learning outcomes involving how test instructions play out in situ, the impact of the teacher in the conversations, and benefits of collegial discussions on assessment.

Research on teachers' CPD acknowledges the futility of assuming a direct causal relationship between training efforts and, for example, student performance resulting from 'better teaching'. Although it has been argued that available models for evaluating the impact of CPD efforts are not comprehensive enough (Day and Sachs 2004b, 27-29), it is reasonable to assume that increased knowledge and awareness have some positive impact, such as how these EFL teachers will prepare their students for the speaking test in the future. The workshop model applied here also meets a number of criteria suggested for successful CPD and research engagement, such as providing case studies relating to teachers' own experiences (cf. Borg 2010, 2015, see also Tinker Sachs and Ho 2011 on using videotaped cases in EFL teacher education).

The possibility for collaborative discussions among teachers is important, since it is a factor that has great impact on teachers' willingness to stay in the profession (Smith and Ingersoll 2004). Teachers also expressed enthusiasm about the potential (and necessity) of collaborative assessment of speaking tests, in line with what the Swedish Schools Inspectorate has suggested. By working together on particular sequences, participants could go back and forth looking at their respective impressions of test-takers' proficiency and scrutinize the grounds for their initial feelings about specific test-takers. By shedding light upon the complexity of perfectly 'normal' and recurrent interactional patterns, participants had the opportunity to see testing talk as a particular form of social interaction where many institutional and interpersonal goals intersect. Even though the workshop method proposed here requires empirical studies and the presence of researchers, we believe there is much to be gained at local schools by using authentic test recordings as grounds for discussions on conducting, understanding, and assessing second/foreign language oral tests. In light of a growing focus on equity in assessments, particularly in standardized tests, the findings presented here reveal that raising EFL teachers' awareness of testing interaction and their own preconceptions about what constitutes 'good' oral proficiency may indeed be a fruitful route toward improved practice.

### **Acknowledgements**

The research presented is funded by the Swedish Research Council [Reg. no. 2012-4129].

### **References**

- Antaki, Charles. 2011. "Applied conversation analysis: from explication to intervention." In *Applied conversation analysis: Intervention and change in institutional talk*, edited by Charles Antaki, 2-14. Basingstoke: Palgrave Macmillan.

- Atkinson, J. Maxwell, and John Heritage, eds. 1984. *Structures of social action*. Cambridge: Cambridge University Press.
- Bolam, Ray, and Agnes McMahon. 2004. "Literature, definitions and models: towards a conceptual map." In *International handbook on the continuing professional development of teachers*, edited by Christopher Day and Judyth Sachs, 3-32. Maidenhead: McGraw Hill Education.
- Borg, Simon. 2010. "Language teacher research engagement." *Language Teaching* 43 (4):391-429. doi: 10.1017/S0261444810000170.
- Borg, Simon. 2015. "Key issues in doing and supporting language teacher research." In *International perspectives on teacher research*, edited by Simon Borg and Hugo Santiago Sanchez, 1-13. Basingstoke: Palgrave Macmillan.
- Cochran-Smith, Marilyn, and Susan L. Lytle. 1999. "Relationships of Knowledge and Practice: Teacher Learning in Communities." *Review of Research in Education* 24 (1):249-305. doi: 10.3102/0091732X024001249.
- Cohen, Louis, Lawrence Manion, and Keith Morrison. 2007. *Research methods in education*. 6 ed. London: Routledge.
- Council of Europe. 2001. *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.
- Day, Christopher, and Judyth Sachs. 2004a. "Professionalism, performativity and empowerment: discourses in the politics, policies and purposes of continuing professional development." In *International handbook on the continuing professional development of teachers*, edited by Christopher Day and Judyth Sachs, 33-63. Maidenhead: McGraw Hill Education.
- Day, Christopher, and Judyth Sachs. 2004b. "Professionalism, performativity and empowerment: discourses in the politics, policies and purposes of continuing professional development." In *International handbook on the continuing professional development of teachers.*, edited by Christopher Day and Judyth Sachs, 33-63. Maidenhead: McGraw Hill Education.
- East, Martin. 2014. "Coming to terms with innovative high-stakes assessment practice: Teachers' viewpoints on assessment reform." *Language Testing*. doi: 10.1177/0265532214544393.
- Ellis, Rod. 1997. *SLA research and language teaching*. Oxford: Oxford University Press.
- Erickson, Gudrun. 2012. National assessment of foreign languages in Sweden. Accessed 14 February, 2014.
- Erlam, Rosemary. 2008. "What do you researchers know about language teaching? Bridging the gap between SLA research and language pedagogy." *Innovation in Language Learning and Teaching* 2 (3):253-267.
- Galaczi, Evelina D. 2008. "Peer-peer interaction in a speaking test: the case of the First Certificate in English examination." *Language Assessment Quarterly* 5 (2):89-119.
- Gustafsson, Jan-Eric, and Gudrun Erickson. 2013. "To trust or not to trust?—teacher marking versus external marking of national tests." *Educational Assessment, Evaluation and Accountability* 25 (1):69-87.
- Hasselgren, Angela. 2000. "The assessment of the English ability of young learners in Norwegian schools: an innovative approach." *Language Testing* 17 (2):261-277.

- Jenkins, Laura, and Markus Reuber. 2014. "Diagnosing seizure patients: A conversation analytic intervention to help neurologists identify diagnostically relevant linguistic features in seizure patients' talk." *Research on Language and Social Interaction* 47 (3):266-279.
- Kasper, Gabriele. 2013. "Managing task uptake in oral proficiency interviews." In *Assessing second language pragmatics*, edited by Steven Ross, J. and Gabriele Kasper, 258-287. Basingstoke: Palgrave Macmillan.
- Kasper, Gabriele, and Steven J. Ross. 2007. "Multiple questions in oral proficiency interviews." *Journal of Pragmatics* 39 (11):2045-2070. doi: 10.1016/j.pragma.2007.07.011.
- Lazaraton, Anne, and Larry Davis. 2008. "A microanalytic perspective on discourse, proficiency, and identity in paired oral assessment." *Language Assessment Quarterly* 4 (4):313-335. doi: 10.1080/15434300802457513.
- Lieberman, Ann. 1996. "Practices that support teacher development: Transforming conceptions of professional learning." In *Teacher learning: New policies, new practices*, edited by M. W. McLaughlin and I. Oberman, 185-201. New York, NY: Columbia University, Teachers' College Press.
- Lundahl, Christian, and S. Tveit. 2014. "Att legitimeras nationella prov i Sverige och i Norge – en fråga om profession och tradition." *Pedagogisk forskning i Sverige* 19 (3-4):297-323.
- Mann, Steve. 2005. "The language teacher's development." *Language Teaching* 38 (3):103-118.
- Markee, Numa P., and Gabriele Kasper. 2004. "Classroom talks: An introduction." *The Modern Language Journal* 88 (4):491-500.
- Messick, Samuel. 1989. "Validity." In *Educational measurement*, edited by Robert L. Linn, 13-103. Washington, DC: American Council on Education/Macmillan.
- Nakatsuhara, Fumiyo. 2008. "Inter-interviewer variation in oral interview tests." *ELT Journal* 62 (3):266-275.
- Nassaji, Hosseim. 2012. "The relationship between SLA research and language pedagogy: Teachers' perspectives." *Language Teaching Research* 16 (3):337-365.
- National Union of Teachers in Sweden. 2006. *Kompetensutveckling bland lärare [Continuous professional development among teachers]*. Stockholm: National Union of Teachers in Sweden.
- Okada, Yusuke, and Tim Greer. 2013. "Pursuing a relevant response in oral proficiency interview role plays." In *Assessing second language pragmatics*, edited by Steven Ross, J. and Gabriele Kasper, 288-310. Basingstoke: Palgrave Macmillan.
- Sandlund, Erica, and Pia Sundqvist. 2011. "Managing task-related trouble in L2 oral proficiency tests: Contrasting interaction data and rater assessment." *Novitas-ROYAL* 5 (1):91-120.
- Sandlund, Erica, and Pia Sundqvist. 2013. "Diverging task orientations in L2 oral proficiency tests – a conversation analytic approach to participant understandings of pre-set discussion tasks." *Nordic Journal of Modern Language Methodology* 2 (1):1-21.
- Sidnell, Jack, and Tanya Stivers, eds. 2013. *The handbook of conversation analysis*. Chichester: Wiley-Blackwell.

- Smith, Thomas M., and Richard M. Ingersoll. 2004. "What are the effects of induction and mentoring on beginning teacher turnover?" *American Educational Research Journal* 41 (3):681–714.
- State Council. 2010. *Ett lyft för den som vill: Utvärdering av den statliga satsningen på fortbildning av lärare, 2010:12 [A boost for the one interested: Evaluation of the national CPD program for teachers, 2010:12]*. Stockholm: State Council.
- Stokoe, Elizabeth. 2011. "Simulated interaction and communication skills training: The 'Conversation Analytic Role-play Method'." In *Applied conversation analysis: Changing institutional practices*, edited by Charles Antaki, 119-139. Basingstoke: Palgrave Macmillan.
- Stokoe, Elizabeth. 2014. "The Conversation Analytic Role-play Method (CARM): A method for training communication skills as an alternative to simulated role-play." *Research on Language and Social Interaction* 47 (3):255-265.
- Sundqvist, Pia. 2009. "Extramural English matters: Out-of-school English and its impact on Swedish ninth graders' oral proficiency and vocabulary." Diss., Karlstad University.
- Swedish Schools Inspectorate. 2012. *Riktad tillsyn av bedömning och betygssättning hos skolor med stora avvikelser vid omrättning av nationella prov [Examination of assessment and grading at schools with great deviations in re-assessments of national tests]*. Stockholm: Swedish Schools Inspectorate.
- Swedish Schools Inspectorate. 2013. *Olikheterna är för stora. Omrättning av nationella prov i grundskolan och gymnasieskolan, 2013 [The differences are too great. Re-assessing national tests in compulsory and upper secondary school, 2013]*. Stockholm: Swedish Schools Inspectorate.
- Tinker Sachs, Gertrude M., and Belinda Ho. 2011. "Using cases in EFL/ESL teacher education." *Innovation in Language Learning and Teaching* 5 (3):273-289.

**Appendix: Transcript notations (adapted from Atkinson and Heritage 1984,****ix-xvi)**

:	Colon(s). Extended/stretched sound, syllable or word.
<u>learn</u>	Emphasis.
(.)	Brief micropause of less than (0.2) seconds
(1.8)	Timed pause
(( <i>scratch</i> ))	Scenic details
.	Falling pitch
?	Rising vocal pitch
↑↓	Marked rising/falling intonation shifts
°word°	Passage of talk noticeably quieter than surrounding talk
[ ]	Overlapping/simultaneous talk
[	Onset of overlap, when overlap completion is difficult to capture
Whe-	Abruptly cut off sound or word
< >	Talk produced noticeably quicker (> <) or slower (< >) than surrounding talk
hhh	Audible aspiration, possibly laughter
.hh	Audible inbreath
Mo(hh)re	Within-speech aspiration, possibly laughter
MINE	Speech noticeably louder than surrounding speech
(pt)	lip smack
\$word\$	talk produced through smile voice
<i>italics</i>	Codeswitching

## Acknowledgements

This work was supported by the Swedish Research Council under Grant 2012-4129.

The authors have no financial interests or benefits arising from the application of this research.

All three authors are researchers in the project, but authorship for this paper is not equally distributed. The first author has assumed the main responsibility for analysis and writing, and was the principal analyst in the study underlying the workshop. The second author has collected the original data and co-authored the article underlying the study. The second author has also participated in designing and conducting workshops. The second author has collaborated closely with the first author in writing the literature review and with the structure and refinement of the manuscript. The third author has collaborated in designing and carrying out workshops, and has participated in discussions on the current study and manuscript revisions.

## About the authors

Erica Sandlund is Associate Professor and senior lecturer of English linguistics at Karlstad University, Sweden. She is a conversation analyst and her research targets social interaction in institutional settings. Current research includes performance appraisal interviews in organizations and second language speaking tests. Sandlund is currently project manager of the Testing Talk project [Swedish Research Council 2012-4129] on the speaking part of the national test of English in Sweden.

Pia Sundqvist is Associate Professor of English linguistics at Karlstad University, Sweden. She worked as a teacher of English, Spanish, and Swedish in secondary and upper secondary school in Sweden for many years before she took up research. Her research interests include second language acquisition, informal learning of English, and assessment of second language oral proficiency.

Lina Nyroos currently holds a postdoctoral research position at the Department of Scandinavian Languages, Uppsala University, Sweden. As a conversation analyst, she studies second language speaking tests as part of the Testing Talk project, and she has also conducted research in other institutional settings, such as academic tutoring sessions and performance appraisal interviews in organizations. She is also a member of Humtank, a national think tank for promoting research in the humanities.