

Diverging task orientations in L2 oral proficiency tests – a conversation analytic approach to participant understandings of pre-set discussion tasks

av Erica Sandlund, Ph. D. & Pia Sundqvist, Ph. D.

Author notes: Erica Sandlund, Senior lecturer, Department of Language, Literature, and Intercultural Studies, Karlstad University, SE-651 88 Karlstad, Sweden, erica.sandlund@kau.se, +46-54-700 1709, Fax +46-54-700 1432 , and Pia Sundqvist, Senior lecturer, Department of Language, Literature, and Intercultural Studies, Karlstad University, SE-651 88 Karlstad, Sweden, pia.sundqvist@kau.se, +46-54-700 1508, Fax +46-54-700 1432.

Correspondence concerning this paper should be addressed to Erica Sandlund, Senior lecturer, Department of Language, Literature, and Intercultural Studies, Karlstad University, SE-651 88 Karlstad, Sweden, erica.sandlund@kau.se, +46-54-700 1709, Fax +46-54-700 1432.

Acknowledgements: The study was made possible through funding from the Center for Language and Literature in Education (CSL) at Karlstad University.

Abstract

Presumably most students strive to do well in school and on national tests. However, even in standardized tests, students' and examiners' expectations on what it means to 'do well' may diverge in ways that are consequential to performance and assessment. In this paper, we examine how students and teachers in an L2 English peer-peer speaking national test (9th grade) display their understandings of appropriate ways of dealing with pre-set discussion tasks. Using conversation analysis and 38 recorded national tests in English in Sweden, we demonstrate, e.g., how teachers' displayed understandings of how tasks should be appropriately handled steer the interactional trajectory between students in particular directions. The analysis shows that participants spend much time on negotiating understandings of the task-at-hand. We argue that in terms of valid assessment of oral proficiency, task understandings merit more attention, as task negotiations inevitably generate different conditions for different dyads and teachers.

1. Introduction

The development and administration of oral proficiency tests with high validity and where students at varying proficiency levels are offered optimal opportunities to demonstrate their skills has been a recurrent topic of interest for linguistic as well as educational researchers (see e.g. Hasselgren 2004, Pollitt and Murray 1996, de Jong and van Ginkel 1992, de Jong and Verhoeven 1992, Lorenzo-Dus 2007, Iwashita et al. 2008). The number of empirical studies taking an interactional approach to oral proficiency tests in a foreign/second language has grown steadily since the 1990s (Young and He 1998, e.g. Lazaraton 1992, Seedhouse and Egbert 2006, Nakatsuhara 2008, Wong and Waring 2010). The explicit focus on the test interaction itself means that the analytic focus lies in identifying and describing how (i.e. through which linguistic, embodied, and social resources) participants (teachers, students, examiners, testees) shape and orient to the ongoing institutional activity. In the present paper, such an approach is taken to paired oral proficiency tests, and we examine how participants treat a pre-set discussion task in situ. Based on data collected from the oral part of the English national test in Sweden, a high-stakes test, we focus on how students and their teacher sometimes show diverging understandings of the task-at-hand, which in turn impacts the topical and interactional trajectory of the oral proficiency test. Our observations underscore not only the centrality of pre-task talk (test instructions), but also offer a methodological route, Conversation Analysis (CA), to uncovering examiners' and learners' preconceived notions of a particular task, displayed in their orientations to the ongoing activity.

CA has become increasingly influential to the study of second language speaking tests (Sacks, Schegloff, and Jefferson 1974, Stivers, Mondada, and Steensig 2011). Thee the analytic process targets the social organization of turns and actions in oral language tests (cf. Lazaraton 1997, 2002, Young and He 1998), and Schegloff, Koshik, Jacoby and Olsher (2002) argue that conversation analytic

research can “inform the design of L2 assessment tasks” and highlight “fundamental issues regarding the positing of appropriate assessment criteria and the interactional processes through which assessment criteria are applied and negotiated” (p. 17).

Since language learning and testing tasks are generally viewed as central to language learning (Hellerman and Pekarek Doehler 2010), one area of interest for interactional approaches to language learning and testing is how tasks are managed by participants in situ; in the classroom as well as in testing contexts (cf. Hellermann 2007, Markee 2005, Sandlund and Sundqvist 2011). Ross (1998, 336) notes that learners’ success in an oral proficiency test is largely dependent on his/her pragmatic competence, which in turn “to no small degree depends on his or her understanding of the form and purpose of questions embedded in the context of the interview”. In a study of adult beginning-level learners negotiating the opening of a new task, Hellermann (2007, 91) noted that “pretask talk is fundamental (...) because it establishes a mutual understanding of the context for their interactions and for rudimentary task planning”. In essence, preparation talk before a task is actually underway is viewed as important for establishing a shared understanding between participants. In a review of studies on learning tasks, Seedhouse (2005) urges researchers to empirically take into account the distinction between a *task-as-workplan*, i.e. the intended pedagogy of a particular task (cf. Breen 1989), and a *task-as-process*, i.e. what actually happens as a task is managed in situ by members. For example, a test task can be planned with certain pedagogical objectives in mind, but once a task (designed *in theory*) is adopted by participants in talk and interaction, their respective understandings and actions may lead to a rather different task trajectory *in practice*. After empirically demonstrating how these two levels diverge in interaction, Seedhouse argues for a shift in focus to the task-as-process level. Similarly, Hellerman and Pekarek Doehler (2010) examined task accomplishment in language classrooms and demonstrated that the same task-as-workplan resulted in rather different task accomplishments by different student dyads.

How, then, do we know that testees (or learners in general) have understood a test task in the way it was originally intended, and in which ways does task understanding appear to matter for speaking test interaction and task accomplishment in situ? Informed by conversation analytic work in the area of language learning and testing tasks, the present paper takes as a point of departure a sequential approach to task accomplishment in a dyadic peer–peer speaking test of L2 English. It should be pointed out that we use L2 as a cover term including also foreign language learning, which is in line with other scholars (see e.g. Ellis and Barkhuizen 2005, Mitchell and Myles 2004). The analysis focuses on how 9th grade students in pairs negotiate a pre-set discussion task in a standardized national test of English, and where their language teachers, who acts as the examiner, also becomes an active participant in the negotiation of task meaning and accomplishment. In Sweden, where the tests were recorded, the national test of English is summative and, therefore, consequential for the testees.

Our main objective is to shed light upon how different orientations to the same test task co-exist and sometimes clash during the course of the speaking test, which in turn impacts how the test task is ultimately performed. An overreaching research interest is instances in dyadic speaking tests where it becomes evident that teachers and students orient to different aspects of the task, and that the diverging approaches themselves become matters for negotiation.

We argue that in terms of valid assessment of students' oral proficiency, displayed orientations to an ongoing test task merit more attention, as the task negotiations inevitably generate different task conditions for different student dyads and teachers. The study also contributes to an enhanced understanding of L2 oral tests as a communicative event where various agendas interplay, and it emphasizes the need for more studies of real-life conditions for students in high-stakes testing contexts.

2. Dyadic oral proficiency testing: The case of the national test of English in Swedish schools

Designing and setting up tests for assessing L2 learners' oral proficiency in the target language is a complex matter and the jury is still out on the best setup for testing L2 performance. One common option is oral proficiency interviews (OPIs), where a testee is alone with an examiner during the test (see e.g. He and Young 1998, Lazaraton 2002, Seedhouse and Egbert 2006). Another setup involves assessing two (or more) learners at the same time (see e.g. Gan, Davison, and Hamp-Lyons 2008). Despite some caution regarding the assessment of individual learners (and an examiner) co-constructing a test interaction, dyadic setups are commonly used in educational settings, mainly because such a format resembles natural conversation (Ducasse and Brown 2009).

The test examined in the present study, the national test of English used in Swedish schools for students in ninth grade, became mandatory in 1998, and is set up as a dyadic peer-peer speaking test. As a national test of a core subject, it is an important part of students' final grade. The test is constructed by members of the so-called NAFS-project (NATIONella prov i Främmande Språk) at the University of Gothenburg. One explicit aim of the test is to aid teachers' summative assessment (Erickson and Börjesson 2001, 257). In addition, the design of the national test allows for teachers to easily integrate it with their teaching. Moreover, the test aligns well with the functional and communicative view of language that is manifest in the curriculum and syllabus. The types of national tests of English used in different countries vary, but the national test of English used in Sweden is a typical so-called proficiency test (see e.g. Brown and Abeywickrama 2010). This means that it does not demand prior knowledge of a particular content; instead the test aims to test learners' general language proficiency. The current national test of English consists of three parts. In part A, focus is on oral interaction and production (i.e. the speaking test), whereas parts B and C test receptive skills (reading and

listening comprehension) and written production (essay). Thorough step-by-step evaluative test-rounds of each part of the test precede every decision about a final inclusion in an official test. Thus, the test is considered both valid and reliable. Test evaluations among English teachers have been very positive; teachers appreciate the fact that the test helps them interpret the syllabus (Erickson and Börjesson 2001) and that it offers a wider repertoire of text genres than what is usually the case at their local schools (Naeslund 2004). For this paper, data is drawn from part A of the national test used in the spring of 2007.

The national test of English has been evaluated on several occasions. For example, on behalf of the Swedish National Agency of Education, Naeslund (2004) evaluated all national subject tests, the national test of English in 9th grade being one of them. Compared with the national tests of Swedish and Mathematics, he found that (i) the English test is rated most favorably and (ii) the tests of English and Swedish are more closely linked to teaching than the one of Mathematics. With regard to the oral parts of the subject tests, Naeslund (2004) concludes that they are dynamic and stimulating for the students but also very demanding, since they require good improvisation skills. The oral tests are demanding also for the teachers, putting particular strain on their attention skills (p. 144). Furthermore, the oral parts of the tests are claimed to be highly relevant in relation to goal attainment.

The English speaking test offers variation with regard to the level of difficulty as well as test design. This variation is important considering the fact that the test is to be taken by students whose English proficiency varies considerably; nevertheless, the test should be stimulating for every single student. Riggenbach (1998) has shown that speaking tests that offer flexibility with regard to conversational topics are more beneficial for interaction than tests that offer fewer options of conversational topics. Supposedly, the interaction between the students tends to become more natural with open or flexible tasks/topics and, as a consequence, the spoken output becomes a better mirror of students' oral proficiency skills. On the other hand, less flexible speaking tasks/topics can be preferred in dyadic speaking tests since such steered tasks/topics might pre-empt that a more skilled student takes over and "ruins" the chances for a less skilled interlocutor on the actual test occasion (Hasselgren 1996).

Recently, the Swedish Schools Inspectorate has carried out a number of large-scale (750 schools) evaluations in which experienced teachers have re-assessed and graded national tests of various subject, including English. In the 2011 report, the deviations between the original test grades and the re-assessed ones were found to be more serious and frequent than expected (The Swedish Schools Inspectorate 2011). As regards the English test, the greatest discrepancy was found for Part C, the essay. It ought to be mentioned that no oral parts were included in the evaluation, most likely because audio data are more difficult to collect than written data. In addition, there are relatively few teachers who actually make audio test recordings – circa 20-25 % for the English test (Gudrun Erickson, personal communication) – despite the fact that the test guidelines clearly encourage teachers to do so. Studies focusing on the oral part of the national tests are,

therefore, particularly timely. As of this writing, the format for the oral part of the national test of English remains the same as the tests from 2007 examined in the present paper. We argue that close analysis of ongoing dyadic test interaction also gives new knowledge about how students approach tasks in standardized speaking tests.

3. Data and method

3.1 Data

The empirical data presented here was originally collected for a study on extramural English (Sundqvist 2009), and encompasses a corpus of 199 recorded speaking tests in English, including the oral part of two old national tests of English, two other similar tests, and finally the oral part of the “real” national test administered in 2007 (“The world around us”). For the purpose of the present analysis, we used conversations from the latter test (38 recorded dyads in total). Since the national test is mandatory and constitutes an important part of the students’ final grade in English, these were the test occasions where the students had most at stake. Furthermore, the national test is designed to align with the curriculum for English in Sweden. Both the curriculum that was in use in 2007 and the present one clearly state the necessity for students to be able to speak and interact in everyday situations in order to receive a pass in English (see The Swedish National Agency of Education 2011). Leaving school with a passing grade in English would correspond to being an Independent user (at approximately level B1) according to the Common European Framework of Reference for Languages (The Swedish National Agency of Education 2012, Council of Europe 2001). The conversations were audio recorded only, since that is what the national guidelines stipulate (i.e., no video). The test conversations have been transcribed in detail according to conversation analytic procedures developed by Gail Jefferson (see e.g. Atkinson and Heritage 1984, Stivers, Mondada, and Steensig 2011). All conversations are dyadic. The students were divided into random pairs by the researcher, but the teacher could change the dyads if, for some reason, the pairings were found to be inappropriate. The students’ own English teacher gave the test instructions and was present throughout the whole test. In our corpus there were three teachers, all with a long experience of teaching secondary school English.

3.2 Method

Our methodological approach is Conversation Analysis (CA). CA is rooted in American sociology and dates back to the late 1960s. Today, CA constitutes a field of inquiry on its own, focusing on both everyday and institutional talk and interaction, with both basic, applied and interventionist research (Antaki 2011). CA uses naturalistic data consisting of recordings (audio and/or video) of talk in

order to understand larger social structures realized in actions in social interaction, such as turn-taking (Sacks, Schegloff, and Jefferson 1974). CA is strongly influenced by the ethnomethodological program of Harold Garfinkel, which emphasizes the everyday methods for meaning-making that people apply as part of everyday practice (Garfinkel 1967), as well as the sociologist Erving Goffman's observations on the systematics of human interaction (e.g. Goffman 1983). More recently, CA has entered the realm of educational research, for example in classroom studies (Seedhouse 2004, Čekaitė 2006, Markee 2005), second language studies (Sandlund and Sundqvist 2011, Hall, Hellerman, and Pekarek Doehler 2011, Mori and Markee 2009, Markee and Seo 2009, Lazaraton 1992, Markee 2000), and in research on learning as social action and practice (Sahlström 2011, Martin 2004, Melander and Sahlström 2010). The analytic process centers on identifying participants' own methods for meaning-making, rather than the analyst's, in recordings of authentic talk, which according to Seedhouse (2005, 257) makes CA "exceptionally strong by comparison to other research methodologies in terms of ecological validity".

The analytic work of CA involves description of systematic features of social interaction as displayed and built up by participants in sequences of talk, which have been transcribed in detail. Collections of phenomena of interest are not usually built on traditional principles of constructing representative samples, "simply because nobody knows what 'representativeness in interaction' means" (Wagner and Gardner 2004, 5), but rather to discern the systematics and internal logic of each sequence through a process of analytic induction (ten Have 1999).

For the present study, we first identified a phenomenon of interest, i.e., displayed understandings of the test task among teachers and students in sequences where these understandings appear to clash and become a local focus of the interaction. Next, we went through our corpus in search of the phenomenon, and this second analytic collection of sequences then underwent additional analysis. In the present paper, we show data from two of these test interactions that in different ways illustrate how task understandings impact task accomplishment in consequential ways. Furthermore, our analytic examples illuminate how close examination of participants' actions and turn-by-turn co-construction of a test task can reveal such understandings, which offers valuable insights into some of the factors that influence test administration and performance. Naturally, with such an analytic approach, we are not arguing that diverging task understandings are problematic in all or even most speaking tests; instead, we demonstrate two cases in which they do seem to overshadow the development of a 'natural-sounding' conversation. This knowledge can be used for raising teacher/examiner awareness of task-related problems and aid the development of test instructions.

4. Managing a discussion task – analysis

Our first fragment (1a-b) is taken from a test conversation between two students, Ann and Mia. Their teacher, who acts as the examiner for the national test, sits in on the test, and the part of the test from which the segment was drawn involves discussion cards (“topic cards”) that the students take turns drawing. Each topic is dealt with until the students themselves (and in some cases, the teacher) brings it to a close, and a new topic card is drawn. The present sequence begins with Ann reading the text on a new topic card (“How important is the way you look and the way you dress?”). She then embarks on her first contribution to the discussion in line 4 (1a):

(1a) “but I think that's stju:pid”

[520522032] TEA = teacher, ANN = female student, MIA = female student

1 ANN ((reads)) how important is: the: way you look an::d the way
2 you dress ((stops reading))
3 (.)
4 ANN (pt) u::h↑ (.) u:h it's- depends (.) on what (1.9) what (0.8) s- (.)
5 style (0.5) eller (0.6) it's ss (.) depends on how you (.)
6 ↑want? to look
7 (0.5)
8 ANN and >I think< I think that u:h (.) you should u:h (.) dress
9 like (1.7) you want to but u:hm (4.1) mmh (1.9) yes (hh)hhhhHUH
10 **TEA but is it impor**tant
11 (3.1)
12 ANN it's- depends on (0.9) hmm (.) hhh yesss

After some displayed hesitation, Ann emphasizes that the matter of looks and appearance is dependent on how a certain person wants to appear. In line 7, there is a short pause that Mia could have utilized as a transition point for a contribution of her own, but instead Ann continues with a personal opinion claim (framed by “I think”) in lines 8-9; i.e., that people should dress according to their personal preferences. Her turn is slightly fragmented and contains several instances of self-repair (see e.g. Schegloff, Jefferson, and Sacks 1977) which indicate that Ann is making an effort to formulate her turn carefully. At this point, Ann has brought forth two aspects of the original topic question; i.e., that it is linked to your personal style preference, and that she thinks that people should dress according to their own preferences.

So far, the conversation is rather typical of the test recordings we have examined. The reader of the topic card should initiate the first discussion turn and then leave room for the other student to continue. However, the phenomenon of interest becomes relevant in line 10, where the teacher, rather unexpectedly, formulates a question in response to Ann’s previous turns: ”but is it impor↓tant”. Her turn is prefaced with the particle “but”, which can be heard as a signal of disagreement, and the question posed emphasizes a particular lexical item in the discussion card topic: ”important”. In this context, immediately succeeding Ann’s second turn, the teacher’s question and emphasis on “important” can be heard as an

objection against the fact that the discussion question has not fully been answered yet. A pause follows after which Ann recycles her initial hesitant “it’s- depends on” (from line 4) in line 12, which signals that Ann insists that the topic card question cannot be answered with a simple *yes* or *no*, despite the fact that the teacher’s question sets up for a yes/no answer. She is displaying some trouble continuing her turn, and when her talk trails off in line 12, Mia, the other student, continues with what can be seen as a proposed continuation of Ann’s turn. However, Mia’s turn also contains an additional aspect on the topic: that the matter of clothes and appearance is linked to the peer group one is part of (lines 13-14, Fragment 1b):

(1b)

13 MIA your friends I think (.) it depends on what (.) type of
 14 friends you have=
 15 ANN yeah,
 16 MIA =if you are in a group (0.6) I think it is very important
 17 (0.4) for someone (0.3) to look (0.5) on a special way (1.5) to
 18 try to fit in?
 19 ANN ye:s?
 20 (2.0)
 21 MIA but I think that's stju:pid (hhhHHh)
 22 ANN yeah (.) u:hm::
 23 (2.6)
 24 TEA any special occasion when you should dress up=
 25 ANN =°yes° whennn (1.7) mm (1.6) when you're going somewhere
 26 (.) mm maybe on (1.4) party or (.) some (.) thing (.) like that
 27 (1.5) can you dress up a little more .hh unusual hhHMH
 28 MIA mm butu:h then you want to:: (0.6) look nice and you think
 29 it's fun to (1.9) dress nice and (2.3) ye:s?
 30 ANN mhm
 31 MIA so (.) but (.) I think it's important to uh (1.9) not think
 32 what everyone else (0.5) >think about you< and try to (1.7)
 33 hhh feel what you want(hhh) (.) yourself (0.5) that is the
 34 best?
 35 ANN °yes°?
 36 ((recording stopped))

In line 16, Mia elaborates on her contribution and states that in certain social groups it can be “important” to fit in by choosing a particular appearance. As we can see, Mia does orient to the teacher’s question by recycling the word *important*, but also to Ann’s repeated efforts to emphasize that the relative importance of clothes and appearance “depends” on a number of social matters. Her turn is finalized with a question intonation, which Ann confirms. After a pause, Mia continues with an evaluative claim: “but I think that's stju:pid (hhhHHh)”. Up to this point, Ann and Mia have brought forth several aspects of the given topic and finalized their contributions with an assessment. A relatively long pause in line 23 follows, after which the teacher adds another question on the topic. This turn may be an attempt to stimulate further dialogue on the topic of clothes, this time steering the discussion toward occasions when people should dress up. Ann takes

the first response turn and formulates, with some hesitation, that parties and similar functions are occasions when the dress code should be different. Her turn is a relatively obvious answer to the teacher's question, but since the question was launched, it prefers an answer from one of the students (cf. Schegloff 2007). However, the teacher's contribution moves the students' discussion from matters of identity, independence, and peer pressure to another possible aspect on the topic, and Mia's response in lines 28-29 indicate that the issue of dressing up for a party was not part of her train of thought. With several characteristics of a so-called dispreferred turn (cf. Pomerantz 1984) such as the initial agreement particle ("mm") and the delay of the actual disagreement (launched with "butu:h then"), Mia's turn brings the topic trajectory back to their preceding talk about choice, peer pressure, and independence, as she emphasizes the free will and enjoyment in dressing up for parties. Her elaboration continues in lines 31-34 with a concluding comment (prefaced by "so", see e.g. Bolden 2006 for an overview), where she yet again recycles the word "important" from the card and the teacher's turn. However, she emphasizes that social pressure to dress in ways that do not match one's own personal preference is not a route to success; instead people should "try and feel" what they are comfortable with. She finalizes her turn with an assessment, "that is the best", where the emphasis on "that" contributes to strengthening that choice and individuality are what she finds central about this topic. Ann agrees, and the recording is then stopped (for reasons unknown to us).

Having examined the sequential development closely, three important observations should be made. Firstly, the students' understanding and approach to the topic seems to be rather similar, in contrast to the teacher's approach, which repeatedly brings up other aspects for discussion. The students' contributions are relatively free from the original wording in the topic card, but the topic is nevertheless dealt with in ways that appear relevant to them. Secondly, the teacher's first contribution does not appear to be occasioned by displays of unsolvable interactional trouble on part of the students – for example, there is no marked pause between Ann's and the teacher's turn in Fragment (1a). The teacher's second contribution (line 24, Fragment 1b), however, comes after a pause, and as we mentioned, we can suspect that the teacher treated the pause as a display of trouble continuing the conversation. On the other hand, a pause of 2.6 seconds is not extremely long in a second language context, and it is entirely possible that the students would have continued elaborating on the thread they had initiated but needed some time to think. Thirdly, it is clear that the teacher's perspective on the topic is different from the students', which affects the topical development. In Fragments (1a-b), the teacher's understanding of how the topic should be treated is displayed in a rather direct orientation toward the written text on the topic card ("How important is the way you look and the way you dress?"). We can only speculate as to why the teacher formulates her turn with such an explicit focus on a lexical item in the question, but her turn can undoubtedly be heard as a request for a fuller answer to the original question. However, Ann is rather persistent in her efforts to reject the question as answerable with agreement or disagreement, and

this approach is also seconded by Mia, even though Mia lexically relates to the degree of “importance” by placing it in relation to peer groups.

The teacher’s second turn does not build on the content of the students’ preceding contributions, but nevertheless, Mia manages to reclaim their focus on independence and freedom from social pressure as she brings her perspective to a close. In sum, the teacher’s displayed understanding of the test task reveals a more literal reading of the text and consequently, a more narrow perspective on how the task should be appropriately handled. By excluding the initial “how” from the topic card, the teacher converts wh-question into a yes/no-question, thereby limiting the scope of relevant answers. In addition, the teacher’s turns were less sequentially linked, which could indicate an understanding of the test task that entails a broad discussion on many aspects of the given topic rather than an elaboration on particular aspects of interest to the students. Also, it is evident that the students in these two segments are rather persistent in pursuing their own approach, even though the teacher’s contributions in some ways obstruct the development of a peer-driven conversation on a topic where they clearly have things to say. As for the test instructions, they only state that the students should try and explain why or why not they agree with a certain proposal or question, and that is, in fact, what Mia and Ann are doing. The analysis, however, shows that the teacher has a slightly different understanding of how this should be accomplished. If teacher and student understandings of an ongoing activity differ markedly, the clash could affect how the teacher then assesses students’ performance in a second language oral proficiency test. To illustrate our analytic point about diverging understandings, we examine a second sequence below.

Fragments (2a-c) involves two ninth graders, Per and Liv. The conversation shows both similarities with and differences from Fragments (1a-b) above, and our transcript begins in lines 38-39 where Liv is reading the text on a new topic card: “What can we learn from history?” The sequence has a rather staggering start as Liv is unsure as to whether the card is referring to history as a school subject or something else, and the teacher replies that it is history in the world that the topic card is after. Liv repeats “in the world” as a confirmation before she embarks on a first commentary in lines 48-49:

(2a) “but what can we learn from history”

[510711072] TEA = teacher, PER = male student, LIV = female student

```
37 (6.6.)
38LIV (pt) `↑happ u::h (.) ((aser)) what can we <le:arn> from
39 history? ((stops reading))
40 (1.2)
41 ((sniffling sound))
42 (4.1)
43LIV °(mm)° (.) history y'know the:: uh (1.5) ämnet1 in the school or
44 (.) .hh
```

¹ Ämnet = the school subject (history)

45 hist[o r y]
46TEA <[no his]tory in the< world (.) [I]mean,
47LIV [ye-]
48 >in the world veh .hh < (1.9) thatu:h the world is going
49 for↑ward (.) for example for four hundred years ago we: had .hhh
50 (1.1) when it was war we:::u hhTHUH HHH (0.6) we didn't (.)
51 shout out guns we:: (.) fight withu:h,
52PER sword
53LIV >yes< (.) and now we- (.) is not good to have a war but .hh now
54 we:: for example shoot with guns ?
55PER HHH[(HH)
56LIV [I (k(hh)no(hh)w?)HHHUH (.) \$and (.) for four hundred
57 years we\$ didn't havu:h (.) cars) and now we have cars? .hhh
58 (1.9)
59PER HHmm?
60 (5.1)

Liv's first comment, "that the world is going forward", is followed by an example about warfare. She is abiding by the test instructions in that she brings up a standpoint, and elaborates on that standpoint with examples, and Per is filling in a word, "sword", that Liv appears to be searching for in line 51. She confirms the proposed word and continues with an evaluative mitigation regarding war as an example ("is not good to have a war but .hh") but returns to the example at hand with "guns" as a sign of world development. With her mitigation, she displays that her example is not grounded in any pro-war approach on her part, but that warfare is nevertheless a relevant example of how things have gone "forward". Per makes a laughter-like sound that is echoed by Liv in line 56 ("I (k(hh)no(hh)w?)HHHUH"). Their laughter may be of the troubles-resistant kind (Jefferson, 1984), indicating a shared understanding that the topic raised is not entirely socially unproblematic, and her "I know" may be an additional sign of this shared understanding. She then continues with a similar example about cars as a means of transportation (i.e. the example also focuses on technological advancements), albeit a less socially problematic one.

A silence as long as 5.1 seconds follows (line 60) and neither Liv nor Per claims the floor. In line 61, the teacher's contribution is probably occasioned by the trailed-off discussion, but note how her turn is packaged:

(2b)
61TEA but what can we ↑learn (.) from history
62LIV thatu:h,
63 (1.2)
64LIV [yea-]
65PER [how] they live?=
66LIV =yeah how they live (.) and that we: are going forward
67TEA °mmm?° (.) can we e- de- (m) (.)u:h (.) (pt) .hhh do you
68 get(hhh) (.) any
69 HELp from knowing that? (0.5) how they lived
70LIV nnn- (.) no? U::h NO?

71 hm?
72LIV >a:h< no? hhUH HUH HUH (HHH) [.HHHH
73TEA [so you think it'sss (0.3) a
74 progress in wa:r and so on,
75 (2.1)
76LIV °nno: (.) (de)°
77 (5.9)

Almost identical to the lexical emphasis on "important" in Fragment (1a), the teacher emphasizes the word "learn", and her turn-initial "but" also indicates some dissatisfaction with prior talk. Just as in (1), it appears as if the teacher is orienting back to the written question, indicating that the topic, so far, has not been treated to satisfaction. Regardless of what the teacher's intentions are, about which we can only speculate, her turn is heard as a type of corrective action where the students are prompted to focus their discussion on what we (people in general) can *learn* from historical events and changes. Liv then makes a few hesitant and cut-off efforts to identify examples of lessons to be learned, which Per overlaps with "how they live" as an example (65). It is possible that the staggering dialogue is colored by the initial confusion on what the topic card meant, as Per's turn brings to mind the usefulness of history knowledge acquired in school rather than to important lessons for humanity from world history in a more abstract sense. Liv immediately catches on and agrees with Per through repetition and then moves on to maintaining her initial claim: that "we are going forward" (66). Even at this point, the students display a partly different notion of the task than what the teacher is attempting to accomplish, as they remain on the theme of development. In lines 67-70, the teacher makes yet another effort to promote the learning aspect of the topic, but reformulated as a fragmented "°mmm?° (.) can we e- de- (m) (.)u:h (.) (pt) .hhh do you get(hhh) (.) any HELp from knowing that? (0.5) how they lived". By reformulating what she is after, using Per's and Liv's own formulation "how they lived", the teacher is attempting to prompt a discussion on whether humans are "helped" or "learn" anything from knowing what it used to be like. Liv responds with disagreement and laughter (line 72). The teacher then attempts a new strategy where she produces a *so*-prefaced upshot of Liv's talk, "so you think it'sss (0.3) a progress in wa:r and so on", a turn shape that seeks confirmation from Liv that this is indeed her opinion. Liv responds hesitantly (76), but her talk trails off and a long pause follows. As a next attempt, the teacher prompts Per to contribute to the topic:

(2c)

78TEA what about Per what do you say.
79PER u:hm
80 (2.9)
81PER .HHHHHHH [(HHHHH)
82TEA [is it good to know about your history.
83 (1.8)
84PER >yeah<

85 (0.9)
86TEA wh(hh)y.
87 (0.6)
88PER uh it's (hhhh)
89 (6.4)
90PER mmmmmm, .hhh
91 (0.4)
92TEA .hh do we learn? From the mistakes we made? .hh
93LIV yes_
94TEA we DO?=
95LIV =hmm
96TEA (k) (.) can you tell- (.)give 'xamples
97 (3.9)
98LIV wellI'm not so good i(HH)n (HH)IS(HH)TOR(HHH)Y (HH)BUT(hh)
99 [(HH)HEH (.) HHEH
100TEA [do you have any examples?
101LIV (hh)huh (.) †hmmmm
102PER >no< (hh)?
103TEA no? (0.7) .HH okay let's take your then
104 ((new topic card introduced))

Per is clearly displaying trouble producing a response and the teacher offers another question prompt requesting an assessment of whether it is “good” to know about one’s history. Having narrowed down Per’s response options and only receiving an affirming “yeah” (line 84) in response, the teacher asks Per to motivate his brief answer: “wh(hh)y.”² When held accountable for justifying his agreement, the same trouble displays and silences appear. The teacher then makes an additional effort to explain what she, in her understanding of the test question, is asking for: “hh do we learn? from the mistakes we made? .hh” (92). Liv agrees, but is then also held accountable for elaborating on her agreement (line 94) and a request for examples (96). Liv’s response this time is of a different kind – she laughingly states that she is not “so good i(HH)n (HH)IS(HH)TOR(HHH)Y”, and yet again, we can see a return to the confusion between world history in general and history in school, and that she treats the teacher’s question as if she is expected to provide historically correct examples of lessons to be learned. Per receives a similar prompt for examples and when he responds negatively, the teacher abandons her attempts to help the students manage the topic in the way that she understood it and she encourages them to draw a new card.

What we can see from the lengthy and almost painfully staggering sequence in Fragments (2a-c) is that the students are displaying repeated difficulties tackling the given topic, and that the teacher is very persistent in trying to steer the discussion toward the topic card and her own understanding of the test task. Just as in Fragment (1a), a component of the wording in the card is emphasized in order to

² It is worth pointing out that since the teacher’s request for an assessment is packaged as a question structurally preferring a positive answer, it would take much more interactional work for Per to respond negatively, regardless of his standpoint in the matter (cf. Pomerantz 1984, Schegloff 2007).

put the spotlight on exactly *what* is to be discussed, but the students nevertheless display difficulties in interpreting and carrying out the discussion according to the teacher's instructions and prompts. Being language teachers ourselves, it is evident that teacher conduct during oral tests like these aims to support and expand students' opportunities to produce assessable output in English, but our analysis also shows something in addition to that: it is not only important that the students talk, but a topic should also be handled in particular ways to pass as acceptable in situ. Possibly, the teacher's repeated attempts to steer Liv away from the warfare topic may be an attempt to divert the focus away from a less politically correct aspect of historical development, but the diverging understandings about how to handle a given topic are visible even after the warfare topic thread has been abandoned. The teacher's contributions can be heard (by the students) as implying additional objectives besides just getting the students to chitchat – there appears to be a pedagogical thread where a deeper critical approach to the task-at-hand is brought forth. The students, however, do not seem to perceive these efforts, and the teacher's repeated attempts to direct the conversation does not result in increased dialogue between Per and Liv (which is the purpose of these peer–peer tests). Rather, her efforts reshape the sequential trajectory to an exchange where the students are prompted to (and do attempt to) answer teacher-driven questions. It is evident that the topic for this sequence was not very suitable for this particular peer dyad, and it affects their opportunities to show what they know rather than what they don't know (cf. Erickson and Börjesson 2001).

5. Implications

Researchers as well as test constructors have spent time on developing valid test instruments that suit as many language learners as possible (see e.g. Erickson and Börjesson 2001, Hasselgren 1996, Lorenzo-Dus 2007). Our study is relevant for both groups but in different ways. From the perspective of interaction research, by analyzing participants' orientations toward the tasks, we have demonstrated that teachers and students display different understandings and expectations. Further, we have demonstrated that these differences affect topic management during the test. In addition, we have shown some other difficulties students experience when their contributions are not accepted as relevant, in particular when students display repeated difficulties in contributing to a certain topic. As was mentioned above, the national test (including its conversation cards) is carefully tried out before a particular test version is officially used, and the purpose of the open and flexible design is that all students should have the opportunity to demonstrate their English oral proficiency skills. However, some students need more structured tasks with explicit instructions (i.e. pretask talk, Hellermann 2007), even though Wong and Waring (2010, 277) in their study of speaking tasks used in language teaching note that “the most authentic interaction often turns out to be off-task talk”. They recommend open tasks since “the designed-ness of tasks (...) goes against the

dynamic nature of interaction to which learners should ideally be exposed” (p. 277). Nevertheless, a downside of open and flexible tasks as the ones examined here may be that the test topics themselves are open to interpretation, which also means that testees and teachers may have perceived the purpose of the task differently. We have demonstrated here that the *task-as-process* in this test was a result of how participants displayed their understandings of the *task-as-workplan* (cf. Breen 1989, Hellerman and Pekarek Doehler 2010, Seedhouse 2005).

Based on our findings, we argue that future teachers would benefit from the inclusion of interaction-based studies in second language teaching methodology courses. The exposure to authentic data in teacher education would raise future teachers’ awareness of what conditions and expectations (institutional as well as interactional) language learners are faced with in the classroom and in speaking tests.

Perhaps the interactional management of diverging task perspectives is most interesting in interactions with role asymmetries (such as learner/teacher). As students are expected to produce gradable output, whereas the teacher has the examiner role, participants have different role-based opportunities for pursuing their own perspectives. As a parallel, Nyroos (2010) examined group tutorials in higher education, focusing on how teachers and students display different understandings of the ongoing activity. While the students to a large degree asked concrete questions about the particular task linked to the group work they were set out to do, the teacher focused on transforming the questions so that they became more abstract, a procedure which suited the teacher’s aim of simultaneously teaching students about scientific methods.

In Nyroos’s analyses there was a clear connection between the teacher’s actions and a pedagogical objective, and the teacher’s actions then become part of doing-being a teacher. In relation to our own analysis, it is reasonable to assume that the teacher’s actions during the speaking test serve the purpose of stimulating the students in order to facilitate their oral interaction. Needless to say, such actions are also part of being an English teacher. In contrast to institutional talk, participants in everyday conversations are usually free to negotiate different perspectives. However, during a national speaking test, the interaction is limited by various institutional restrictions, such as task goals, roles, and time frames. For instance, the asymmetry of the teacher–student relationship becomes salient when the teacher asks questions: the students have to relate to them somehow, or account for why they choose not to. In addition, the teacher is about to assess and grade his or her students’ oral production and has, accordingly, epistemic primacy (cf. Stivers, Mondada, and Steensig 2011, 13) regarding proper test performance. By way of example, we have shown that the students indeed hear and treat the teacher’s turns as questions they are accountable for answering, rather than as prompts for stimulating the conversation. In turn, the conversation takes the form of an interview (cf. Seedhouse and Egbert 2006) between the students and the teacher rather than between the students themselves.

Needless to say, we can only speculate on the development of these particular conversational sequences had the teacher remained passive, and our aim has not

been to evaluate teacher conduct *per se*. However, our observations align well with those of a classroom study focusing on planned interactions between Americans learning Japanese and Japanese natives (Mori 2002). In the Mori study, the classroom interactions did not result in conversations that resembled natural conversations; instead, the interactions turned into interviews with questions and searches for suitable responses. In our analysis, when the teacher highlights one aspect of a given topic as more important than another, it is possible to view the teacher's actions as counterproductive, at least in relation to the explicit aim of the test and the fact that oral fluency is viewed as more important than content (cf. Naeslund 2004). Our analyses pinpoint the importance of preparing students for oral tests by explicitly discussing test instructions and expectations in the classroom prior to the test occasion. Albeit small, our study clearly shows that teachers and students orient to the tasks at hand very differently. The teachers' displayed actions in our fragments can be partly explained by an objective to elicit as much assessable talk as possible from their students, partly by other reasons, such as helping out when a student fails to reflect on and contribute to topics that relate to the fundamental values and tasks of the school (cf. Fragment 2c). Separating oral fluency and other cognitive abilities and knowledge during the actual test, then, presents a challenge to teachers, and for students, the test requires more than mere oral production. Thus, participation in the national speaking test also demands some degree of knowledge of the topics that are to be discussed, something which we have addressed in a previous study (Sandlund and Sundqvist 2011).

6. Concluding remarks

In our article, we have focused on how different interactional treatments of a given test task has consequences for the test talk trajectory. Furthermore, we have demonstrated a methodological pathway for identifying participants' (teachers' and students') task understandings. We argue that careful attention to sequential structures in oral test interactions also has the potential to inform testing and assessment practice (see also Sandlund and Sundqvist 2011, 2012). In line with this, we have initiated a research project (*Testing Talk*, <http://www.kau.se/testing-talk>) where, among other things, English teachers take part in practice-oriented training with authentic test recordings and are given the opportunity to reflect on test procedures and conditions for learner participation in oral tests.

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.
- Bolden, Galina B. 2006. "Little words that matter: Discourse markers "so" and "oh" and the doing of other-attentiveness in social interaction." *Journal of Communication* no. 56 (4):661-688.
- Breen, Michael P. 1989. "The evaluation cycle for language learning tasks." In *The second language curriculum*, edited by Robert K. Johnson, 187-206. Cambridge: Cambridge University Press.
- Brown, H. Douglas, and Priyanvada Abeywickrama. 2010. *Language assessment. Principles and classroom practices*. 2 ed. White Plains, NY: Pearson Education.
- Čekaitė, Asta. 2006. *Getting started: Children's participation and language learning in an L2 classroom*, Linköping University, Linköping.
- Council of Europe. 2001. *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.
- de Jong, John H.A.L., and Lieneke W. van Ginkel. 1992. "Dimensions in oral foreign language proficiency." In *The construct of language proficiency: Applications of psychological models to language assessment*, edited by Ludo Verhoeven and John H.A.L. de Jong, 187-205. Amsterdam: John Benjamins.
- de Jong, John H.A.L., and Ludo Verhoeven. 1992. "Modeling and assessing language proficiency." In *The construct of language proficiency: Applications of psychological models to language assessment*, edited by Ludo Verhoeven and John H.A.L. de Jong, 3-19. Amsterdam: John Benjamins.
- Ducasse, Ana Maria, and Annie Brown. 2009. "Assessing paired orals: Raters' orientation to interaction." *Language Testing* no. 26 (3):423-443.
- Ellis, Rod, and Gary Barkhuizen. 2005. *Analysing learner language*. Oxford: Oxford University Press.
- Erickson, Gudrun, and Lena Börjesson. 2001. "Bedömning av språkfärdighet i nationella prov och bedömningsmaterial." In *Språkboken*, edited by Rolf Ferm and Per Malmberg, 255-269. Stockholm: Myndigheten för skolutveckling.
- Gan, Zhengdong, Chris Davison, and Liz Hamp-Lyons. 2008. "Topic negotiation in peer group oral assessment situations: A conversation analytic approach." *Applied Linguistics* no. 30 (3):315-344.
- Garfinkel, H. 1967. *Studies in ethnomethodology*. Englewood Cliffs: Prentice Hall.
- Goffman, Erving. 1983. "The interaction order." *American Sociological Review* no. 48 (1):1-17.
- Hall, Joan Kelly, John Hellerman, and Simona Pekarek Doehler, eds. 2011. *L2 interactional competence and development*. Edited by David Singleton, *Second Language Acquisition*. Bristol: Multilingual Matters.

- Hasselgren, Angela. 1996. "The EVA Project - diagnostic tests in English for the 8th class, or something more?" *Språk og språkundervisning* (2):27-31.
- Hasselgren, Angela. 2004. *Testing the spoken English of young Norwegians: A study of test validity and the role of 'smallwords' in contributing to pupils' fluency*. Edited by Michael Milanovic and Cyril J. Weir, *Studies in Language Testing*. Cambridge: Cambridge University Press.
- He, Agnes Weiyun, and Richard Young. 1998. "Language proficiency interviews: A discourse approach." In *Talking and testing: Discourse approaches to the assessment of oral proficiency*, edited by Richard Young and Agnes Weiyun He, 1-24. Amsterdam: John Benjamins.
- Hellerman, John, and Simona Pekarek Doehler. 2010. "On the contingent nature of language-learning tasks." *Classroom Discourse* no. 1 (1):25-45.
- Hellermann, John. 2007. "The development of practices for action in classroom dyadic interaction: Focus on task openings." *The Modern Language Journal* no. 91 (1):83-96.
- Iwashita, Noriko, Annie Brown, Tim McNamara, and Sally O'Hagan. 2008. "Assessed levels of second language speaking proficiency: How distinct?" *Applied Linguistics* no. 29 (1):24-49.
- Lazaraton, Anne. 1992. "The structural organization of a language interview: A conversation analytic perspective." *System* no. 20 (3):373-386.
- Lazaraton, Anne. 1997. "Preference organization in oral proficiency interviews: the case of language ability assessments." *Research on Language & Social Interaction* no. 30 (1):53-72.
- Lazaraton, Anne. 2002. *A qualitative approach to the validation of oral language tests*. Cambridge: Cambridge University Press.
- Lorenzo-Dus, Nuria. 2007. "The best of both worlds? Combined methodological approaches to the assessment of vocabulary in oral proficiency interviews." In *Modelling and assessing vocabulary knowledge*, edited by Helmut Daller, James Milton and Jeanine Treffers-Daller, 220-233. Cambridge: Cambridge University Press.
- Markee, Numa. 2000. *Conversation analysis*. Mahwah, NJ: Lawrence Erlbaum.
- Markee, Numa. 2005. "The organization of off-task talk in second language classrooms." In *Applying conversation analysis*, edited by Keith Richards and Paul Seedhouse, 197-213. Basingstoke: Palgrave Macmillan.
- Markee, Numa, and Mi-suk Seo. 2009. "Learning talk analysis." *IRAL* no. 47 (1): 37-63.
- Martin, Cathrin. 2004. *From other to self. On learning as interactional change*, Humanistisk-samhällsvetenskapliga vetenskapsområdet, Uppsala University, Uppsala.
- Melander, Helen, and Fritjof Sahlström. 2010. *Lärande i interaktion*. Stockholm: Liber.
- Mitchell, Rosamond, and Florence Myles. 2004. *Second language learning theories*. 2 ed. London: Arnold.

- Mori, Junko. 2002. "Task design, plan, and development of talk-in-interaction: A select study of small group activity in a Japanese classroom." *Applied Linguistics* no. 23 (3):323-347.
- Mori, Junko, and Numa Markee. 2009. "Language learning, cognition, and interactional practices: An introduction." *IRAL* no. 47 (1):1-9.
- Naeslund, Lars. 2004. *Prövostenar i praktiken. Grundskolans nationella provsystem i ljuset av användares synpunkter*. Stockholm: Skolverket [The Swedish National Agency of Education].
- Nakatsuhara, Fumiyo. 2008. "Inter-interviewer variation in oral interview tests." *ELT Journal* no. 62 (3):266-275.
- Nyroos, Lina. 2010. "'Om du har något annat ord för det?' Hur deltagares skilda perspektiv kommer till uttryck i ett undervisningssammanhang." *Språk & Stil* no. 20:165-188.
- Pollitt, Alastair, and Neil L. Murray. 1996. "What raters *really* pay attention to." In *Performance testing, cognition and assessment. Selected papers from the 15th Language Testing Research Colloquium, Cambridge and Arnhem*, edited by Michael Milanovic and Nick Saville, 74-91. Cambridge: Cambridge University Press.
- Pomerantz, Anita. 1984. "Agreeing and disagreeing with assessments: some features of preferred/dispreferred turn shapes." In *Structures of social action*, edited by J. Maxwell Atkinson and John Heritage, 57-101. Cambridge: Cambridge University Press.
- Riggenbach, Heidi. 1998. "Evaluating learner interaction skills: Conversation at the micro level." In *Talking and testing: Discourse approaches to the assessment of oral proficiency*, edited by Richard Young and Agnes Weiyun He, 53-67. Amsterdam: John Benjamins.
- Ross, Steven. 1998. "Divergent frame interpretations in language proficiency interview interaction." In *Talking and testing: Discourse approaches to the assessment of oral proficiency*, edited by Richard Young and Agnes Weiyun He, 333-353. Amsterdam: John Benjamins.
- Sacks, Harvey, Emmanuel A. Schegloff, and Gail Jefferson. 1974. "A simplest systematics for the organization of turn-taking in conversation." *Language* no. 50 (4):696-735.
- Sahlström, Fritjof. 2011. "Learning as social action." In *L2 interactional development and competence*, edited by Joan Kelly Hall, John Hellerman and Simona Pekarek Doehler, 45-65. Bristol: Multilingual Matters.
- Sandlund, Erica, and Pia Sundqvist. 2011. "Managing task-related trouble in L2 oral proficiency tests: Contrasting interaction data and rater assessment." *Novitas-ROYAL (Research on Youth and Language)* no. 5 (1):91-120.
- Sandlund, Erica, and Pia Sundqvist. 2012. "'But what can we learn from history' – om konsten att diskutera 'rätt' under det muntliga nationella provet i engelska." *KAPET* no. 8 (1):45-66.
- Schegloff, Emmanuel A. 2007. *Sequence organization in interaction. A primer in conversation analysis*. Cambridge: Cambridge University Press.

- Schegloff, Emmanuel A., Gail Jefferson, and Harvey Sacks. 1977. "The preference for self-correction in the organization of repair in conversation." *Language* no. 53 (1-2):361-382.
- Schegloff, Emmanuel A., Irene Koshik, Sally Jacoby, and David Olsher. 2002. "Conversation analysis and applied linguistics." *Annual Review of Applied Linguistics* no. 22:3-31.
- Seedhouse, Paul. 2004. *The interactional architecture of the language classroom: A conversation analysis perspective*. Oxford: Blackwell.
- Seedhouse, Paul. 2005. "'Task' as research construct." *Language Learning* no. 55 (3):533-570.
- Seedhouse, Paul, and Maria Egbert. 2006. The interactional organisation of the IELTS speaking test. In *IELTS Research Reports*, http://www.ielts.org/pdf/Vol6_Report6.pdf.
- Stivers, T, L Mondada, and J Steensig, eds. 2011. *Knowledge and morality in conversation. Rights, responsibilities and accountability*. Cambridge: Cambridge University Press.
- Sundqvist, Pia. 2009. *Extramural English matters: Out-of-school English and its impact on Swedish ninth graders' oral proficiency and vocabulary*, Faculty of Arts and Education - English, Karlstad University, Karlstad.
- ten Have, P. 1999. *Doing conversation analysis*. London: Sage Publications.
- The Swedish National Agency of Education. 2011. *Curriculum for the compulsory school, preschool class and leisure-time centre 2011*. Stockholm: The Swedish National Agency of Education.
- The Swedish National Agency of Education. 2012. *Internationella språkstudien 2011. Rapport 375*. Stockholm: The Swedish National Agency of Education.
- The Swedish Schools Inspectorate. 2011. *Engelska i grundskolans årskurser 6-9. Kvalitetsgranskning. Rapport 2011:7*. Stockholm: The Swedish Schools Inspectorate.
- Wagner, Johannes, and Rod Gardner. 2004. "Introduction." In *Second language conversations*, edited by Rod Gardner and Johannes Wagner, 1-17. London: Continuum.
- Wong, Jean, and Hansun Zhang Waring. 2010. *Conversation analysis and second language pedagogy*. New York, NY: Routledge.
- Young, Richard, and Agnes Weiyun He, eds. 1998. *Talking and testing: Discourse approaches to the assessment of oral proficiency*, *Studies in Bilingualism* 14. Amsterdam: John Benjamins