

Running Head: Interactional patterns in test performance

**International engineering graduate students' interactional patterns
on a paired speaking test: Interlocutors' perspectives**

Corresponding author: Talia Isaacs

Graduate School of Education

University of Bristol

35 Berkeley Square, Clifton

Bristol, United Kingdom BS8 1JA

E-mail: talia.isaacs@bristol.ac.uk

Isaacs, T. (2013). International engineering graduate students' interactional patterns on a paired speaking test: Interlocutors' perspectives. In K. McDonough & A. Mackey (Eds.), *Second language interaction in diverse educational settings* (pp. 227–246). Amsterdam: John Benjamins.

Abstract

This study examines interactional patterns between English language learners from different first language backgrounds on a collaborative speaking task from the Business English Certificates, a standardized test often used to screen prospective employees for their English language proficiency at international companies. Peer interactions of 84 international engineering graduate students, categorized using Galaczi's (2008) interactional typology, were examined in relation to individual interlocutors' target language use in daily life, oral proficiency level, and perceptions of their performance on the task. Results showed that collaborative pairs were the most positive overall about the quality of their interactions. Conversely, dominant members of unevenly matched (asymmetric) pairs rated their interactional success and their ability to understand and be understood by their partner most negatively.

Word length: 120

Introduction

Canada welcomes hundreds of thousands of immigrants each year, approximately 80% of whom speak neither official language (English or French) as their mother tongue (Corbeil & Blaser, 2007). For many, the need to acquire a new language compounds the challenge of integrating into society and securing employment. Although the government finances language training courses to assist newcomers in overcoming language barriers and transitioning into the labor market, intercultural communication can pose difficulties in workplace settings. For example, “culturally different” job candidates’ aims to obtain employment may be derailed at the interview stage by linguistic and/or nonlinguistic factors (Laroche & Rutherford, 2007). Even when newcomers are recruited into jobs commensurate with their professional training, oral communication challenges may be prevalent, and those related to intelligibility, broadly defined as second language (L2) learners’ ability to understand and be understood by their interlocutors (e.g., colleagues, clients), appear to be among the most pressing (Derwing & Munro, 2009). For instance, Canadian-born employees often find it difficult to communicate with their foreign-born colleagues and may avoid interacting with them, with oral language barriers serving as a major deterrent (Derwing & Munro, 2008). In companies that emphasize teamwork and collaboration, communication difficulties could impede employee motivation and productivity.

Problems with intercultural communication are not unique to corporate settings and extend to other contexts, including educational settings. The academic domain is of particular importance in Canada due to the influx of “visa students” on Canadian campuses (Belkhodja, 2011) and the need for accessible language support in cases when the medium of instruction is different than the student’s first language (L1). Engaging in social interactions with members of the host culture has been identified as a major adjustment challenge for international students and may have negative repercussions on intercultural adaptation, academic attainment, and social cohesion, particularly in cases where there is

little mixing with members of different cultural or linguistic communities (Spencer-Oatey & Xiong, 2006; Zimmerman, 1995). In addition to requiring oral communication skills to perform their academic tasks, some international graduate students bear research or instructional responsibilities and, therefore, may need to communicate effectively with undergraduate students in addition to their professors and peers. Pronunciation is often identified as a prominent problem by multiple stakeholders including, in the case of international teaching assistants (ITAs), for example, L2 teaching and testing professionals, undergraduate students, and ITAs themselves (Cheng, Myles, & Curtis, 2004; Hoekje & Williams, 1994). Although L2 accents tend to be accorded disproportionate attention due to their perceptual salience, even occasionally serving as a scapegoat to veil discriminatory attitudes (Derwing & Munro, 2009), in some cases, genuine pronunciation problems impede listener understanding (Isaacs, 2008).

The literature on language and communication challenges in workplace and academic settings is relevant to the present study, which examines the nature of the interactions that arise among paired international engineering graduate students from different linguistic and cultural backgrounds on a collaborative L2 speaking test task. This target population arguably need spoken English to perform their academic responsibilities, including, in some cases, instructional duties, to achieve their stated aim of securing employment in North America post-graduation, and to perform domain-relevant tasks, including interacting with various stakeholders (e.g., coworkers, clients) in their eventual jobs (Human Resources and Skills Development Canada, 2011). The study also explores the link between the students' reported L2 use and proficiency, the overall interactional pattern of their co-constructed discourse, and their perceptions of the quality of the communicative exchange and of the role of pronunciation in shaping their interactions. Before describing the language needs and background characteristics of this target population, the remainder of the literature review will overview research on the paired speaking test format (i.e., peer interactions among L2 test-takers), including both language

testing and SLA research on interactional patterns and interlocutor proficiency effects on dyadic interactions. This research is relevant to the present study, which makes use of a collaborative speaking task on a standardized international L2 proficiency test—the Business English Certificates (BEC).

The paired speaking test format and insights from SLA interaction research

There is a growing trend in L2 assessment research to examine the paired speaking test format, as is evidenced by the publication of a recent special issue in *Language Testing* on peer interactions in L2 assessment settings (Taylor & Wigglesworth, 2009). Ducasse and Brown (2009) chart a concomitant movement away from the traditional oral proficiency interview in operational assessment settings, which has been the most common and preferred method for assessing L2 speech since the communicative era (Luoma, 2004), to an increased focus on peer performance on interactional tasks (see Winke, this volume). This has come about as a result of influence from SLA research in support of the Interaction Hypothesis (Long, 1996), which holds interaction is beneficial for L2 learning, with several meta-analyses demonstrating the facilitative effects of interaction (e.g., mediated by corrective or interactional feedback) on the acquisition of target language forms (Li, 2010; Lyster & Saito, 2010; Mackey & Goo, 2007). The increasing prominence of peer oral proficiency testing has been buttressed by claims that pair and group work in language classrooms reflect good teaching practice, promote positive washback, and have the advantage of not containing the power imbalance between interviewer and interviewee that is inherent in the traditional oral proficiency interview (Fulcher, 1996; Winke, this volume), thereby enabling students to perform their best on the assessment (Egyud & Glover, 2001; Taylor, 2001). Further, in comparison with semi-direct (i.e., machine-mediated) nonreciprocal tasks, which involve the test-taker speaking into a recording device without a human audience (e.g., the TOEFL iBT speaking component), face-to-face interactions may contribute to more authentic assessments (i.e., if they

resemble the collaborative tasks that test-takers are likely to perform in the real-world context being generalized to) and tend to appeal more to test-takers (Qian, 2009).

The paired speaking test format was first adopted in the Cambridge ESOL Main Suite Exams in 1996 in response to advancements in L2 research and classroom practice coupled with the need to draw on a broader range of tasks to elicit different dimensions of the L2 oral proficiency construct (Saville & Hargreaves, 1999). Collaborative tasks remain a staple of these tests today, including the BEC, which is the source of the task in the present study. However, test-taker pairings at testing sites tend to be ad hoc, with little attention paid to interlocutor variables that could affect the nature of the interaction and unduly optimize or penalize test-taker performance. That is, involving a conversational partner in L2 oral assessments introduces a range of peer interlocutor variables (e.g., age, gender, personality characteristics, attitudinal variables, L1 background, L2 proficiency level, exposure to the interlocutor's accent) that are extraneous to the construct being measured and that could pose a threat to the validity of the assessment (Van Moere, 2006). To mitigate concerns about test fairness due to haphazard peer pairing practices in operational assessment settings, different interactional patterns are elicited in the Cambridge Main Suite Exams. In addition to the collaborative task, candidates also respond to short interview questions from the examiner and perform a brief monologue. The resulting range of tasks offers opportunities for the use of more varied language than would be possible if all oral proficiency tasks were monologic or in oral interview format (Saville & Hargreaves, 1999).

One of the most extensively researched interlocutor variables in both the language testing and SLA literature is the effect of peer proficiency on L2 learner performance and interactional patterns (e.g., Davis, 2009; Kim & McDonough, 2008), since this research has direct implications for optimizing peer pairing practices in both high-stakes assessment and lower-stakes classroom contexts. In language assessment research, the major focus has been on interlocutor proficiency effects on test-takers'

attainment, as attested by raters' holistic or analytic scoring of learners' paired performances on collaborative L2 speaking test tasks (Davis, 2009; Norton, 2005). Conversely, the orientation in the SLA interaction literature has been on learners' attention to language in collaborative discourse, as attested by the occurrence and resolution of language-related episodes when performing tasks designed to gauge learners' attention to form in L2 classroom settings (e.g., dictogloss; Kowal & Swain, 1994). Overall, findings from the language assessment literature have been inconclusive with respect to rated outcomes, with some studies revealing that test-takers tend to be rated more highly when paired with a higher proficiency interlocutor (Iwashita, 1996), while others reveal no interlocutor proficiency effects on rated measures (Csépes, 2002). However, the quantity of the oral discourse the dyad generates (e.g., total words produced, number of speaker turns) appears to be sensitive to differences in test-taker L2 proficiency level across assessment studies, with test-takers producing relatively more output when paired with a higher-level peer than with a lower-level peer (Davis, 2009). Similarly, findings from SLA research have revealed that language-related episodes tend to occur and to be resolved with greater frequency in pairings with a relatively more advanced L2 interlocutor than with a lower proficiency partner (e.g., Kim & McDonough, 2008; Leeser, 2004).

Recent language assessment and SLA interaction studies have additionally examined interlocutor proficiency in conjunction with the pattern of interaction that emerges between learner dyads on a collaborative task. Storch's (2002) typology of co-constructed discourse, which distinguishes four patterns of interaction in learner dyads' communicative exchanges, has been used to classify interactional discourse in classroom talk in the SLA literature (e.g., Watanabe & Swain, 2007). It has also served as the basis of Galaczi's (2004, 2008) slightly adapted framework for analyzing collaborative speaking test data which, in turn, has been used in language assessment studies on the paired speaking test format (e.g., May, 2009). In both Storch's (2002) and Galaczi's (2008) frameworks,

categorization decisions about the nature of the interactions are made on the basis of the notions of equality (the extent to which the interlocutors contribute proportionally to and assume ownership of the task) and mutuality (the extent to which the interlocutors are engaged with each other's contributions and actively co-construct meaning). Taken together, studies which have employed either framework have shown that interactions which are collaborative in nature (i.e., involve the interlocutors' active interchange of ideas and engagement) foster greater opportunities for noticing form-meaning relationships (Kim & McDonough, 2008; Watanabe & Swain, 2007), are overall rated highly in terms of individual speakers' proficiency scores (Davis, 2009), and pose fewer dilemmas for raters' scoring decisions than dyads in which one interlocutor engages in interactional dominance (May, 2009).

In the language assessment literature, interactional patterns classified using Galaczi's (2008) typology have been investigated with respect to raters' perceptions of the paired interaction and views on whether test-taker performances should be subject to individual or joint scoring (May, 2009). However, the relationship between interactional patterns and *interlocutors'* (i.e., L2 test-takers') perceptions of the communicative efficiency of the interactions and attitudes toward engaging with their partner have yet to be examined. Further, in reference to the engineering graduate student participants in this study, because international students' interactions with members of the target language community have been linked to academic attainment and because oral communication difficulties can act as a deterrent to learners' engagement in L2 oral interactions and as a barrier to social integration (Cheng et al., 2004; Zimmerman, 1995), there is a need to examine whether interactional patterns on a collaborative L2 speaking test task relate to learners' use of and exposure to the L2 in the host country. Therefore, the goal of this descriptive study is threefold: to examine (1) the nature of the interactions that occur between international graduate students on a paired collaborative L2 speaking test task, (2) whether the identified interactional pattern for the dyads relates to individual interlocutors' reported L2

use and proficiency level, and (3) individual interlocutors' perceptions of the quality of their interaction (e.g., in terms of communicative efficiency) in relation to the identified interactional pattern.

Method

Description of context and participants' language needs and interactional practice

This study took place in the oil-rich province of Alberta, Canada, which has suffered from an acute shortage of engineers over the past several decades. Although considerable government resources have been allocated to attracting and retaining foreign engineers, the oral communication challenges that they face are considerable. The “Essential Skills Inventory” (Human Resources and Skills Development Canada, 2011), developed in reference to the Canadian Language Benchmarks, suggests that engineers require L2 oral proficiency skills of up to ‘4’ out of ‘5’ on a language complexity scale, including interacting with numerous stakeholders (e.g., co-workers, clients, suppliers) to carry out their professional tasks. Studies conducted in the Alberta context specifically have demonstrated the detrimental effects of foreign-born engineers’ pragmatically inappropriate responses on job recruiters’ hiring decisions (Louw, Derwing, and Abbott, 2010) and Canadian-born engineers’ reported difficulties communicating with their foreign-born colleagues and avoidance of conversational exchanges with them (Derwing & Munro, 2008). In addition to jeopardizing workplace and company productivity, oral communication challenges can lead to cultural segregation and a lack of social cohesion.

Fuelled by the belief that holding a Canadian degree makes them more employable than their foreign credentials, many engineers seeking to build a career in Canada pursue postgraduate degrees at higher education institutions in hopes of securing a job in industry or academia. This is the case for the 84 international graduate student engineers who participated in this study (57 males, 27 females; $M_{age} = 26.75$; 23–36), enrolled in electrical (49), chemical (32), civil (2), and mechanical engineering programs

(1) at a research-intensive Alberta university. The participants, who all reported having normal hearing, were from numerous L1 backgrounds, including Mandarin (23), Farsi (23), Urdu (9), Bengali (8), Hindi (6), Spanish (4), Sinhala (3), Punjabi (2), and one L1 speaker of each of Cantonese, French, Hindko, Konkani, Pashto, and Portuguese. They were all “visa students” who had moved to Canada on average 1.5 years earlier to pursue graduate studies (range: 1 month – 5 years, 5 months) with the exception of one participant, who had moved to Canada as an undergraduate student seven years earlier and had subsequently obtained citizenship. At least 12 research participants had been admitted to the university without any oral proficiency testing due to having taken versions of the TOEFL that did not include a speaking component (i.e., the paper-and-pencil or computer-based TOEFL). There was no additional language screening for the 40 participants employed as ITAs, including 21 with full instructional responsibilities (e.g., formal lecturing or running labs and tutorials). In addition, none of the participants had taken an English for Academic Purposes course at the university, which some reported was inaccessible due to the expense.

Over 75% of all graduate students at the Faculty of Engineering were foreign nationals at the time of admission, with the largest groups from Mainland China and Iran. Thus, in cases of interactions between peers or professors who did not share the same (or a mutually intelligible) L1, English was used as the lingua franca. Participants estimated interacting in English 44% of the time in their daily lives ($SD = 24$) and 64% of the time at university ($SD = 25$), with participants from Mandarin and Farsi L1 groups, the two largest L1 groups in the dataset and at the faculty, below average in their interactional engagement in English at university (61% and 51% of the time, respectively). Although the large standard deviations (25 for each) evidence considerable variability, with some participants pursuing opportunities to engage in L2 interactions more actively than others, patterns emerged in semi-structured interviews conducted at the end of the data collection session (see description of procedure below). As

one L1 Mandarin speaker related, “My supervisor is Chinese and he recruited student all from Chinese. All from China. So we, our laboratory, are all Chinese. Yeah, whole laboratory. So sometimes we do not have much chance to speak English because we all speak Mandarin.” Like several of his peers, he revealed that his Chinese supervisor was the only one he ever conversed with in English due to her belief that English was professionally important. Similarly, several Iranian participants reported having no opportunities to speak English since moving to Canada. One Iranian student, for example, reported speaking English less in Canada than during the few hours a week she had worked as an IELTS trainer in Iran, due to peer pressure among Iranian students in the faculty to conform by speaking their L1 (although a few Farsi speakers resisted this). She related, “everyone [Iranian students] know that they had the better English, before they come here.” Still other participants reported that their interview with the researcher was the first time they had ever conversed with a Canadian-born individual besides routine encounters. On the basis of participants’ descriptions, a picture of a linguistically and culturally-segregated graduate student body emerges, particularly for students in predetermined cliques who come to the university specifically to work with a supervisor from their country and who are designated for work in a linguistically homogeneous lab from the outset. Challenges in cross-cultural communication emerged in anecdotes. For example, an L1 Bengali speaker described his experience attempting to get help solving an engineering problem from a classmate:

After the class, I passed one of my classmate. He was from China. Firstly he told me can you please write it down, what you are asking. After the class, I asked him, can you please explain me how it [the problem] worked? He told me to write down in a paper what my question is. After writing the question, he told me ok, I know how it works, but I cannot explain it to you in English.

Clearly, the written medium was being used to compensate for poor oral/aural English skills or for reticence to communicatively engage with his peer. In a second anecdote, the participant described larger-scale compensatory strategies for understanding a professor:

One instructor was a native Canadian. So, his English was very fluent and it was a little different. It was very fast and his accent was a little different than what we normally hear. So I could not follow some part of his lecture. So we had some uh Chinese classmate and they could not follow what he taught because he was speaking so fast. The next day of our class he [the teacher] found all the Chinese students, they brought the Chinese version of the textbook.

Apparently, none of the students in the class had asked the instructor to slow down or repeat. It is possible that the instructor was unaware of his students' comprehension difficulties (mostly international students who had recently arrived in Canada) and made little effort to adapt his speech. Regardless, these accounts reveal cross-cultural communication challenges at the faculty due to language barriers.

Instruments and data collection procedure

Because the majority of graduate students at the Faculty of Engineering are L2 English speakers and there is evidence of linguistic segregation, the intention here was to examine the interactions of dyads from different L1 backgrounds. To create pairings that were as linguistically and culturally different as possible, speakers from Indic languages (e.g., Urdu, Bengali, Hindi) and Romance languages (e.g., Spanish, Portuguese, French) were not paired with a peer from within the same L1 group. After completing a language background questionnaire, paired participants completed a series of L2 speaking tasks. Performance on the BEC collaborative task will be the focus of this manuscript. Cambridge ESOL's BEC is a standardized test used by major international engineering companies (e.g., Ericsson, Shell, Vodafone) to screen prospective employees for English language proficiency or for staff development (<http://www.cambridgeesol.org>). The sample collaborative speaking task for the "Vantage Level" test, aligned with level B2 of the Common European Framework of Reference, was used to elicit the peer dialogue. The two interlocutors were each provided with the written scenario of needing to make preparations for a business trip in a foreign country, with discussion points on making travel and accommodation arrangements and learning about foreign business customs. In accordance with exam

procedures, participants were instructed to discuss and decide about these points together (University of Cambridge ESOL Examinations, 2008).

Speech samples were recorded in a quiet room using a multi-track digital recorder, with unidirectional lavalier microphones provided for each interlocutor. To simulate the BEC test procedure, which includes the presence of two examiners, the researcher and a research assistant (RA) were both present during data collection and undertook different roles. The researcher introduced the speaking tasks, operated the recording equipment, interjected if necessary to keep participants on track (e.g., to redress the communication balance if one participant was dominant, in accordance with guidelines described in the BEC handbook), and monitored time. The researcher was not a trained Cambridge rater but had served as examiner for speaking assessments for university entrance purposes and used the Cambridge ESOL standard oral scripts. The RA fulfilled the role of silent observer, recording details on body language and other features that emerged during the interaction in a research log. Although in an operational testing situation, the BEC examiner is instructed to stop the test-taker pair after the approximate 2 minute duration of the task if it does not come to a natural close, for the purposes of data collection in the present study, the time for task completion was not constrained beyond the 2 minute minimum and, thus, was longer and more variable ($M = 4 \text{ min } 42 \text{ s}$; $SD = 2 \text{ min } 5 \text{ s}$). After completing the speaking tasks, both participants filled out separate follow-up questionnaires items on the success of their interactions, how well they had understood and been understood by their partner, the role of pronunciation in their communicative exchange, and their willingness to engage with their partner on future professional projects on separate 5-point Likert-type scales. Finally, the researcher and RA interviewed the interlocutors in separate rooms about their impressions of the interaction with their partner, their oral communication needs and interactional engagement at the university, in their

envisioned future job, and in life more generally. Participants were remunerated for approximately 1.5 hours of their time.

Data analysis

After normalizing interlocutors' speech samples on separate audio tracks, the L2 speech data were transcribed using conversational analysis conventions adapted from Atkinson and Heritage (1984; see Appendix). Next, three categories from Galaczi's (2004, 2008) typology of interactional patterns, derived from an analysis of the topical moves on the collaborative speaking task from another Cambridge test (the First Certificate in English), were used to classify the overall interactional orientation of the paired graduate student engineers' co-constructed discourse on the BEC. Galaczi's first category, the "collaborative" interaction type, which is equivalent to the "collaborative" pattern in Storch's (2002) framework for classroom work, is characterized by both interlocutors contributing creatively to the task and evidence of mutual engagement through peer topic extension and, occasionally, the completion of each other's thoughts. An example of collaborative discourse is shown in Hassan (L1 Iranian male) and Sanjay's (L1 Hindi male) 35 second conversational exchange in

Excerpt 1. Pseudonyms are used for all participants in this paper.

Excerpt 1. Collaborative interaction

- H: where we stay should be close to (.) I'd rather- (.) it be li- (.) close to the, to the (.) to the m- to the meeting to the business mee[ting-],
S: [yeah] we sh- we should look for [that]
- H: [bus]iness place=
S: =yeah we should be close to the place and yeah. Traveling,
- H: that's easy. that's (.) google map.=
S: =e- yeah that's google maps.
- H: yeah.
S: and we have smart phones most of the people. [so]
H: [yeah]
- S: company has given us that. ((laughter))=
- H: =yeah. we shou- we should buy a [company, um],
S: [yeah we should] companie[s:]
H: [we] should ask the

S: company to buy us uh (.) new google ph[ones],
 S: [new] google pha. [yeah].
 H: [is it] android?=

S: =an[droid].
 H: [or] droid? what is [it?] [is it] android?=

S: [android]. yeah.

Notably, both interlocutors propose novel content (Hassan: location of meeting; use of Google maps; Sanjay: company provision of smart phones for employees) and play off of each other's ideas. Their speech is characterized by frequent acknowledgment tokens to signal agreement (particularly in Sanjay's case) and numerous instances of overlapping or latched speech, resulting in little (if any) time between turns.

Galaczi's (2004) second category of interactional behavior, "parallel," is characterized by both interlocutors initiating novel ideas but not following up on each other's turns. This interactional pattern is construed somewhat differently from the "dominant/dominant" category in Storch's (2002) framework. In Galaczi's (2004) conception of the parallel interaction type and its application in the present study, the key notion is not conversational dominance or the inability of the pair to achieve consensus, as is emphasized in Storch's (2002) framework, but, rather, that there is little evidence of mutual engagement (Galaczi, 2004), with few instances of supportive topic development or of the partners' ideas intertwining. For example, in Ayaz (L1 Urdu male) and Alda's (L1 Spanish female) 42 second conversational exchange in Excerpt 2, both parties independently contribute ideas but do not build on their partner's contributions and do not appear to be responsive to the content of the partner's utterances.

Excerpt 2. Parallel interaction

Ay: so we need to know how- many persons for how many person we are (.) booking, in a h- in a hotel? or any other place? (.6) so: uh tickets and: accommodations, and (.) uh, what else? should we: (.) you know take (.6) to that trip (.) with us,

- (1.7)
- Al: um: if they have some: um: special necessities? like um (.7) if some of them cannot eat certain type of food, if they are (.) some like- one of them could be in wheelchairs, and that kind of (1.0) [um: (.) different]
- Ay: [ok th- important] documents? (1.2) and um (...) the documents we will be (1.0) needing in (.7) for meeting?

Ayaz's invitation to Alda to suggest items to bring on the trip is followed by Alda's input on the seemingly unrelated topic of accommodating special needs. Instead of continuing in her line of thought, Ayaz declines to expand her topic and, instead, advances his own idea about needing important documents. Although, in this pairing, there does not appear to be competition for holding the floor, each partner's focus is clearly on getting his/her own points across. Notably, both participants' contributions are marked by dysfluencies at awkward junctures (i.e., within clauses).

Galaczi's third interactional category, "asymmetric," which subsumes Storch's (2002) "expert/novice" and "dominant/passive" interactional patterns, is characterized by one interlocutor clearly dominating the conversation in terms of the amount of discourse produced and the contribution of content, and the other partner assuming a relatively passive role. These unequal contributions may be manifested in the dominant partner's reticence to cede the floor, in the passive partner's lack of initiative in assuming the floor and/or proposing new ideas, or in the dominant partner's collegial scaffolding of the passive partner to facilitate his/her participation in the interaction. Due to the imbalance inherent in this category, the individual interlocutors who made up the asymmetric pairings in this study were further classified as "dominant speaker" and "passive speaker." Excerpt 3 features a 29 second asymmetric conversational exchange between Faisal (L1 Urdu male), who assumed a dominant role in the interaction, and Walton (L1 Mandarin male), who assumed a passive role.

Excerpt 3. Asymmetric interaction

- F: so I will say that uh I have a Canadian boss and I am going with him to Pakistan and he's taking me to Pakistan because I am a Pakistani? [and] I know (.) about th=

W: [yeah]
 W: =kay
 F: their culture a:nd how things go over there?
 W: okay.
 F: so: (.) the first uh (.) thing we have to discuss like what travel and accommodation arrangements we have to make before the trip so:=
 W: =uh huh= [uh huh]
 F: =the first and the most important thing is the ticket (.) [arrange]ment because=
 W: [uh huh]
 W: =yup=
 F: =usually [uh] there are some peak seasons?
 W: [yeah]
 W: okay,
 F: in June and in November?

Faisal made it difficult for Walton to interject, with hardly any pauses between his turns, despite occasionally ending tone groups with rising intonation. He also unilaterally decided on a foreign country in which the BEC scenario would take place and chose a cultural context for which Walton presumably had no cultural reference. This arguably strengthened his position of power in the interaction as the purveyor of knowledge. Walton's contribution was relegated to backchannelling and he was unable to assert control over the task.

Galaczi (2004) refers to a final “blended” interaction type, which is a hybrid of two categories, presented in alteration within the same pairing. This category was not used in the present study because it was possible to categorize the overall interactional orientation of the dyadic discourse using one of the three main interactional patterns described above. In rare instances when the interactional orientation appeared to shift midway through the task and the predominant orientation was unclear, the interactional pattern that was most representative of the first minute of the conversation was selected.

Following initial classification of the data, an intercoder reliability check by a second RA yielded exact agreement on 88% of the category assignments for 60% of the data. In instances of disputes, a third coder (the author) independently recoded the data and made the final classification decision. Next,

descriptive statistics were calculated for participant self-report data, grouped by interactional category, on their English language ability, use, and the quality of their interactions.

Results

Table 1 shows the number of pairs classified into the three interaction types based on their co-constructed discourse, with interlocutors in the asymmetric group further designated as dominant or passive. This breakdown is similar to Galaczi's (2008) study, where the number of collaborative versus parallel pairings was equal, with relatively fewer asymmetric pairings. Indicative patterns suggest a link between interaction type and participants' pre-task questionnaire responses on their language proficiency and use. First, the findings support the claim that asymmetric interactional patterns generally arise in pairings with uneven L2 proficiency profiles (May, 2009). The dominant speakers in this study reported considerably higher English language proficiency, speaking and listening to English a higher proportion of the time, and engaging more with native English speakers than did their passive partners. Notably, these disparities in target language use did not extend to their interactions with L2 speakers from other L1 backgrounds. There is likely a linguistic and cultural component to this finding that needs to be unpacked. For example, although Mandarin speakers accounted for 27% of participants in this study, they were represented in 58% of the cases of individuals classified as passive within the asymmetric group. Similarly, speakers of Indic languages accounted for 27% of participants in this study but were overrepresented in the asymmetric dominant group at 83%. In one such asymmetric pairing, a dominant male L1 Hindko speaker produced 417 words 2.5 minutes into the task compared to his passive female L1 Mandarin speaking partner's 34 words. When the researcher intervened to ask the passive speaker about business customs in an attempt to redress the conversation, the passive speaker reflected on cultural differences in conversational norms, stating, "I think here [in Canada], people can

speak in the meeting but in China, only the head while he's speaking no one can speak, so you have to respect the leader." Thus, interlocutors' views of the pragmatic acceptability of interrupting a partner's turns through overlapping speech and their pre-interactional perceptions of their partner as equal or unequal in status (i.e., through first impressions) are potential influences on turn-taking behavior that warrant further investigation.

Table 1. Interlocutors' reported English language use and proficiency grouped by interaction type

Mean participant self-report data (<i>SD</i>) obtained prior to administering the L2 speech tasks						
Interaction type	No. of pairs (<i>n</i> = 42) & proportion	English speaking & listening proficiency ^a	English spoken in daily life ^b	English spoken at university ^b	Time speaking to native speakers ^b	Time speaking to L2 learners from other L1 backgrounds ^b
<u>Collaborative</u>	<u>15</u> = 36%	<u>3.8 (.6)</u>	<u>45.7 (27)</u>	<u>69.5 (25)</u>	<u>44.7 (23)</u>	<u>32.0 (22)</u>
<u>Parallel</u>	<u>16</u> = 38%	<u>3.7 (.7)</u>	<u>34.3 (19)</u>	<u>56.5 (26)</u>	<u>44.3 (28)</u>	<u>33.0 (25)</u>
<u>Asymmetric</u>	<u>11</u> = 26%	<u>3.9 (.6)</u>	<u>52.3 (22)</u>	<u>67.2 (23)</u>	<u>52.5 (26)</u>	<u>34.2 (22)</u>
<i>Dominant</i>		<u>4.3 (.5)</u>	<u>60.4 (24)</u>	<u>82.1 (14)</u>	<u>62.5 (23)</u>	<u>34.2 (26)</u>
<i>Passive</i>		<u>3.5 (.3)</u>	<u>44.2 (19)</u>	<u>52.5 (22)</u>	<u>34.2 (19)</u>	<u>34.2 (19)</u>

Note: ^aMeasured on a 5-point scale (1 = *extremely poor*, 5 = *extremely proficient*). ^bMeasured on a 0–100% scale.

Taken together, the highest reported English language proficiency and use for dominant speakers followed by collaborative, parallel, then passive speakers suggests that interactional type categorizations could efficiently predict individual speakers' L2 use and proficiency. However, May (2009) found that speakers classified as dominant based on their interaction with a lower proficiency learner may be categorized as collaborative when paired with an interlocutor of similar or higher L2 proficiency (see also Kim & McDonough, 2008). Thus, the interactional pattern clearly depends on the *pairing* and not simply on the individual speaker.

Table 2, which presents descriptive statistics of the interlocutors' perceptions of the interaction obtained in a follow-up questionnaire after all L2 speaking tasks had been completed, also shows clear patterns. Participants whose discourse was collaborative had the most positive overall feeling about the quality of the interaction, reported the fewest gaps in understanding, felt that pronunciation detracted the least from their communicative exchanges, and were the most receptive to working with their partner on future professional projects. The parallel group scored less well on these metrics, albeit better than the asymmetric group, who was the least satisfied with their interactional experience.

Table 2. Interlocutors' perceptions of their interaction and partner grouped by interaction type

Interactional type (n pairs)	Mean participant self-report data (<i>SD</i>) obtained immediately after the L2 speech tasks had been completed				
	Overall success of the interaction	How well understand partner	How well understood by partner	Pronunciation an issue in the interaction	How well would collaborate with partner on a professional task
<u>Collaborative</u> (15)	<u>4.3 (.6)</u>	<u>4.5 (.7)</u>	<u>4.1 (.5)</u>	<u>2.6 (1.3)</u>	<u>4.2 (.7)</u>
<u>Parallel</u> (16)	<u>4.1 (.7)</u>	<u>4.2 (.9)</u>	<u>3.9 (.8)</u>	<u>2.9 (1.2)</u>	<u>3.8 (.7)</u>
<u>Asymmetric</u> (11)	<u>3.8 (.8)</u>	<u>3.8 (1.0)</u>	<u>3.9 (.8)</u>	<u>3.0 (1.1)</u>	<u>3.8 (.6)</u>
<i>Dominant</i>	3.5 (1.0)	3.4 (1.1)	3.8 (.6)	2.8 (1.4)	3.7 (.5)
<i>Passive</i>	4.1 (.5)	4.1 (.8)	4.1 (.9)	3.3 (.9)	3.9 (.7)

Note: ^aMeasured on a 5-point scale (1 = negative response/no, 5 = positive response/yes).

Breaking down the asymmetric group further, the passive speakers were overall more positive than the dominant speakers about the success of the interaction. They also perceived higher mutual understanding and were more open to future collaborations with their partner, recording approximately the same levels of satisfaction as the parallel group. The dominant speakers appear to have been more frustrated with the interaction. However, the passive speakers identified pronunciation as being more problematic than did the dominant speakers, although the high standard deviation for the dominant speakers reveals considerable variability on this point. It could be that the passive speakers were overall

more self-conscious about their pronunciation due to everyday communication difficulties that they experienced, even though they felt supported by their partner in the research setting and were possibly unaware of their partner's more negative perceptions.

Engaging in a collaborative interactional pattern thus appears to be the most positive for L2 learners and could lead to positive rapport-building and establishing solidarity that is important in human relationships (e.g., for job hiring decisions; Kerekes, 2006). Parallel interaction, where there is little uptake on partners' turns, appears to be more preferable, from the interlocutor's perspective, to being paired with someone who struggles to contribute to the conversation and who plays a subordinate role in defining the creative direction of the task. In high-stakes assessments (e.g., operational administrations of the BEC), test-taker pairing practices could affect the quality of the co-constructed discourse and could, in turn, impact raters' scoring decisions. This issue is discussed further in the Discussion section.

Discussion

The goal of this descriptive study was, first, to examine the nature of the interactions that arise between international engineering graduate student dyads from different L1 backgrounds on the BEC collaborative speaking test task. Analysis of the co-constructed discourse generated by each of the pairs revealed that parallel and collaborative interaction types were the most prevalent in the dataset and occurred in approximately equal proportion (together, 3/4 of the time), whereas the asymmetric pattern, which occurred in the remainder of the cases, was less common. This closely corresponds with the proportion of interaction types observed in Galaczi (2008), although, in the present study, it was possible to reliably classify the overall interactional orientation of the dyadic discourse without resorting to a "blended" category (i.e., consisting of two distinct interaction types).

The second goal of the study was to examine individual interlocutors' estimates of their L2 speaking ability and the proportion of their time spent engaging in target language interactions in relation to the interaction type to which they had been assigned. Mean L2 speaking proficiency and language use measures, pooled over interaction type, were the highest for dominant speakers followed by collaborative and parallel speakers, with dominant speakers' passive partners reporting the lowest overall means. This finding supports evidence from the language assessment literature that asymmetric interactions tend to occur when there is a mismatch in paired test-takers' L2 proficiency level (May, 2009) and echoes indications from SLA classroom interaction research that learner pairs with largely discrepant L2 proficiency levels are less likely to be collaboratively oriented in their discourse productions than when they are more evenly matched (Kowal & Swain, 1994). Cultural issues that mediate interactional orientations and turn-taking strategies arose in some participants' qualitative accounts in the present study. These need to be systematically investigated in future research, particularly in relation to other interlocutor variables such as gender and interlocutors' perceptions of gender roles.

The final goal of the study was to examine individual interlocutors' perceptions of their joint performance (e.g., communicative efficiency) in relation to interactional type. Results revealed that interlocutors in the collaborative group were the most satisfied of all groups with the quality of their interaction, experienced the fewest pronunciation and communication problems, and were most receptive to the idea of participating in future professional collaborations with their partner. This suggests that collaborative interactions, which, in previous studies, were found to yield the highest ratings in the paired speaking test format (Davis, 2009) and to be the most facilitative for creating opportunities for learning in classroom settings (Kim & McDonough, 2008), also tend to be perceived the most positively by L2 interlocutors themselves. Further, the finding implies that being in a

collaborative interactional situation with a peer is not only optimal in assessment and classroom settings from the perspective of external parties observing or evaluating the interactions (e.g., raters scoring their performance or researchers examining language-related episodes), but is also conducive to performance from the perspective of the L2 interlocutors who are directly involved in the interaction.

In comparison with the collaborative dyads, interlocutors classified in the parallel and asymmetric-passive groups were relatively less positive about the interactional outcome and their partner, likely due to poorer peer engagement (e.g., as signaled through a lack of topic extension). However, the asymmetric-dominant group appeared to experience the most frustration with the interaction, scoring the interactional indices and their ability to understand and be understood by their partner most negatively, although they ascribed a less important role to pronunciation in contributing to the difficulties than did their passive partners. The passive speakers may have been more concerned about their own pronunciation due to oral communication difficulties that they routinely experience. and, therefore, to have weighted this factor more heavily than the speakers in the other interactional groups, although this is merely speculative. In sum, being a dominant speaker within an asymmetric paring appears to be the worst case scenario from the perspective of an interlocutor performing a collaborative L2 speaking test task. This aligns with May's (2009) contention that raters have the most difficulty arriving at scoring decisions for asymmetric pairings due to the imbalance inherent in the interaction. The issue of whether co-constructed learner performance on an interactional task should be singly or jointly scored, which has been the subject of considerable attention in the language assessment literature (Taylor & Wigglesworth, 2009), was not examined here, since ratings based on the dyadic BEC performances were not obtained. Future research could elicit the ratings and perceptions of interviewers employed at an engineering company with a culturally diverse workforce, for example.

Notably, no objective measure of L2 proficiency was obtained in the present study. Reliance on participant self-report ratings of their L2 oral proficiency, exposure to the target language, and performance on the task, as was done here, has obvious limitations which need to be acknowledged. A further limitation is that, due to the use of descriptive rather than inferential statistics in the present study to examine interlocutors' reported L2 proficiency level, interactional encounters, and perceptions of their interaction, no group differences as a function of interaction type can be ascertained. These issues notwithstanding, the interlocutors' self-report indices yielded clear patterns when grouped by interaction type, to the extent that it was possible to rank order interactional groups on all indicators in a way that was consistent with what has been reported previously in the literature (e.g., that collaborative interactions are viewed most favorably from learning and assessment standpoints; Davis, 2009; Storch, 2002). Further, the value of test-taker self-assessments of their performance has been emphasized in the language assessment literature as an important part of learner awareness-raising (see Alderson, 2005; Winke, this volume). Therefore, the relationship between interlocutors' perceptions of their contribution to the joint performance on the paired-speaking test format, including whether they viewed themselves or their partner to be interactionally dominant, the scores that they obtain, and the feedback that they receive about their performance in score reporting forms needs to be examined in future research.

Although Swain suggests that peer speaking tests may be a way of "biasing for best" (2001, p. 298) in language assessments, asymmetric pairings appear to pose difficulties for key stakeholders in the assessment process, including the dominant speakers in this study, and raters, as shown in previous L2 assessment research (May, 2009). Evidently, peer pairing could have the potential to alter the nature of the interaction (Kim & McDonough, 2008) and to penalize scoring (Iwashita, 1996). Although engineering test-taker pairings in operational assessment situations may be logistically difficult to implement, the relationship between interactional type, L2 proficiency, target language interactions, and

interlocutors' (test-takers') perceptions of their performance needs to be more extensively researched in the interest of test fairness and in order to allay test-takers' concerns. To do this, multiple sources of evidence need to be examined in research settings in order to investigate the extent to which interlocutors' discourse productions, self-ratings of their joint performance, and interview comments about their interaction align. Although the present study represents a preliminary attempt to examine test-takers' perspectives and features a relatively large sample size (42 pairs), one of the limitations is that interlocutors' qualitative interview data and the analysis of their collaborative discourse were not directly mapped onto their quantitative self-report data. These data sources need to be more rigorously triangulated.

The population of engineering professionals in this study, including 1/4 who reportedly bore full instructional responsibilities as ITAs, were motivated to perform well academically and to integrate into the Canadian labor market post-graduation. Clearly, this population require L2 oral communication skills to successfully integrate into academic or workplace culture and to carry out their professional responsibilities. The use of the BEC in this study is ecologically valid inasmuch as it is used for screening or professional development purposes at major international engineering companies. However, the content of the task (organizing a business trip for a company) arguably does not simulate the domain-specific tasks that the international engineers in this study are likely to encounter in academic or workplace settings. Achieving greater authenticity in task design would likely entail enlisting the help of a domain expert (engineer) in the design of a joint problem-solving task, for example (Douglas, 2000), although the implementation of such a task in the present study would have been challenging due to pairings between engineers from different fields (e.g., electrical and chemical engineering). Therefore, the generalizability of the performances on the collaborative BEC speaking task

used in this study to oral communication tasks that engineers are expected to perform in real-world contexts (Human Resources and Skills Development Canada, 2011) is arguably limited.

One of the major challenges that international students on foreign campuses face is engaging in interactions with members of the host community (Cheng et al., 2004; Spencer-Oatey & Xiong, 2006). Due to the high concentration of L2 graduate students and staff at the Faculty of Engineering from which the participants in this study were recruited, opportunities to interact in English in the academic domain are more likely to occur with other international engineers than with Canadian-born individuals. Thus, being paired with an international graduate student peer in this study reflects the grouping that might take place in the university setting if learners move beyond being entrenched in L1 cliques to engaging more broadly with L2 speakers from different linguistic and cultural backgrounds.

References

- Alderson, J. C. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment*. London: Continuum.
- Atkinson, J. M., & Heritage, J. (Eds.). (1984). *Structures of social action*. Cambridge: Cambridge University Press.
- Belkhodja, C. (Ed.). (2011). International migration: The emergence of the mobile student. *Canadian Diversity*, 8.
- Cheng, L., Myles, J., & Curtis, A. (2004). Targeting language support for non-native English-speaking graduate students at a Canadian university. *TESL Canada Journal*, 21, 50–71.
- Corbeil, J. P., & Blaser, C. (2007). *The evolving linguistic portrait, 2006 census*. Ottawa, ON: Statistics Canada.
- Csépes, I. (2002). Is testing speaking in pairs disadvantageous for students? A quantitative study of partner effects on oral test scores. *novELTy*, 9, 22–45.
- Davis, L. (2009). The influence of interlocutor proficiency in a paired oral assessment. *Language Testing*, 26, 367–396.
- Derwing, T. M., & Munro, M. J. (2008, June). *Engineering better communication in the workplace*. Paper presented at the Canadian Association of Applied Linguistics conference, Vancouver, BC.
- Derwing, T. M., & Munro, M. J. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching*, 42, 1–15.
- Douglas, D. (2000). *Assessing languages for specific purposes*. Cambridge: Cambridge University Press.
- Ducasse, A., & Brown, A. (2009). Assessing paired orals: Raters' orientation to interaction. *Language Testing*, 26, 423–443.

- Egyud, G., & Glover, P. (2001). Oral testing in pairs: A secondary school perspective. *ELT Journal*, 55, 70–76.
- Fulcher, G. (1996). Testing tasks: Issues in task design and the group oral. *Language Testing*, 13, 23–51.
- Galaczi, E. D. (2004). *Peer-peer interaction in a paired speaking test: The case of the First Certificate in English*. Unpublished doctoral dissertation, Columbia University, New York.
- Galaczi, E. D. (2008). Peer-peer interaction in a speaking test: The case of the First Certificate in English examination. *Language Assessment Quarterly*, 5, 89–119.
- Hoekje, B., & Williams, J. (1994). Communicative competence as a theoretical framework for ITA education. In C. G. Madden & C. L. Myers (Eds.), *Discourse and performance of international teaching assistants* (pp. 11–26). Alexandria, VA: TESOL.
- Human Resources and Skills Development Canada. (2011). *Essential skills*. Retrieved July 6, 2011, from http://www10.hrsdc.gc.ca/es/english/all_profiles.aspx
- Isaacs, T. (2008). Towards defining a valid assessment criterion of pronunciation proficiency in non-native English speaking graduate students. *Canadian Modern Language Review*, 64, 555–580.
- Iwashita, N. (1996). The validity of the paired interview format in oral performance testing. *Melbourne Papers in Language Testing*, 5, 51–65.
- Kerekes, J. A. (2006). Winning an interviewer's trust in a gatekeeping encounter. *Language in Society*, 35, 27–57.
- Kim, Y., & McDonough, K. (2008). The effect of interlocutor proficiency on the collaborative dialogue between Korean as a second language learners. *Language Teaching Research*, 12, 211–234.
- Kowal, M., & Swain, M. (1994). Using collaborative language production tasks to promote students' language awareness. *Language Awareness*, 3, 73–93.

- Laroche, L., & Rutherford, D. (2007). *Recruiting, retaining, and promoting culturally different employees*. Oxford: Oxford University Press.
- Leeser, M. J. (2004). Learner proficiency and focus on form during collaborative dialogue. *Language Teaching Research*, 8, 55–81.
- Li, S. (2010). The effectiveness of corrective feedback in SLA: A meta-analysis. *Language Learning*, 60, 309–365.
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of language acquisition*. San Diego, CA: Academic Press.
- Louw, K. J., Derwing, T. M., & Abbott, M. L. (2010). Teaching pragmatics to L2 learners for the workplace: The job interview. *Canadian Modern Language Review*, 66, 739–758.
- Luoma, S. (2004). *Assessing speaking*. Cambridge: Cambridge University Press.
- Lyster, R., & Saito, K. (2010). Oral feedback in classroom SLA. *Studies in Second Language Acquisition*, 32, 265–302.
- Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 407–451). Oxford: Oxford University Press.
- May, L. (2009). Co-constructed interaction in a paired speaking test: The rater's perspective. *Language Testing*, 26, 397–421.
- Norton, J. (2005). The paired format in the Cambridge Speaking Tests. *ELT Journal*, 59, 287–297.
- Qian, D. D. (2009). Comparing direct and semi-direct modes for speaking assessment: Affective effects on test takers. *Language Assessment Quarterly*, 6, 113–125.
- Saville, N., & Hargreaves, P. (1999). Assessing speaking in the revised FCE. *ELT Journal*, 53, 42–51.

- Spencer-Oatey, H., & Xiong, Z. (2006). Chinese students' psychological and sociocultural adjustments to Britain: An empirical study. *Language, Culture and Curriculum, 19*, 37–53.
- Storch, N. (2002). Patterns of interaction in ESL pair work. *Language Learning, 52*, 119–158.
- Swain, M. (2001). Examining dialogue: Another approach to content specification and to validating inferences drawn from test scores. *Language Testing, 18*, 275–302.
- Taylor, L. (2001). The paired speaking test format: Recent studies. *University of Cambridge Local Examinations Syndicate Research Notes, 6*, 15–17.
- Taylor, L., & Wigglesworth, G. (Eds.). (2009). Pair work in L2 assessment context [Special issue]. *Language Testing, 26*.
- University of Cambridge ESOL Examinations. (2008). *Business English Certificates: Handbook for teachers*. Cambridge: UCLES.
- Van Moere, A. (2006). Validity evidence in a university group oral test. *Language Testing, 23*, 411–440.
- Watanabe, Y., & Swain, M. (2007). Effects of proficiency differences and patterns of pair interaction on second language learning: collaborative dialogue between adult ESL learners. *Language Teaching Research, 11*, 121–142.
- Winke, P. (in press). The effectiveness of interactive group orals for placement testing. In K. McDonough & A. Mackey (Eds.), *Interaction in diverse educational settings*. Amsterdam: John Benjamins.
- Zimmerman, S. (1995). Perceptions of intercultural communication competence and international student adaptation to an American campus. *Communication Education, 44*, 321–335.

Appendix: Transcription conventions

[] overlapping utterance

= latched utterance, no intervening time between turns

? rising intonation

. falling intonation

, level intonation

- abruptly ended sound

(.) brief silent pause < .5 seconds

(.5) length of pause if \geq .5 seconds

: lengthening of sound

(()) nonlinguistic occurrence (e.g., laughter)

(?) inaudible/incomprehensible utterance

Adapted from Atkinson & Heritage (1984)