

# **EXPLORING THE POTENTIAL FOR ASSESSING INTERACTIONAL AND PRAGMATIC COMPETENCE IN SEMI-DIRECT SPEAKING TESTS**

VS/2021/001

**Fumiyo Nakatsuhara**, CRELLA, University of Bedfordshire, UK

**Lyn May**, Queensland University of Technology, Australia

**Chihiro Inoue**, CRELLA, University of Bedfordshire, UK

**Edit Willcox-Ficzere**, Oxford Brookes University, UK

**Carolyn Westbrook**, British Council, UK

**Richard Spiby**, British Council, UK

BRITISH COUNCIL VALIDATION SERIES

SERIES EDITORS: RICHARD SPIBY AND CAROLYN WESTBROOK

ISSN 2398-7979

© BRITISH COUNCIL 2021

# ABSTRACT

---

To explore the potential of a semi-direct speaking test to assess a wider range of communicative language ability, the researchers developed four semi-direct speaking tasks – two designed to elicit features of interactional competence (IC) and two designed to elicit features of pragmatic competence (PC). The four tasks, as well as one benchmarking task, were piloted with 48 test-takers in China and Austria whose proficiency ranged from CEFR B1 to C. A post-test feedback survey was administered to all test-takers, after which selected test-takers were interviewed.

A total of 184 task performances were analysed to identify interactional moves utilised by test-takers across three proficiency groups (i.e., B1, B2 and C). Data indicated that test-takers at higher levels employed a wider variety of interactional moves. They made use of concurring concessions and counter views when seeking to persuade a (hypothetical) conversational partner to change opinions in the IC tasks, and they projected upcoming requests and made face-related statements in the PC tasks, seemingly to pre-empt a conversational partner's negative response to the request.

The test-takers perceived the tasks to be highly authentic and found the video input useful in understanding the target audience of simulated interactions.

## Keywords

interactional competence, pragmatic competence, task design, semi-direct speaking test, test-taker perceptions, speaking test construct

## Authors

---

**Fumiyo Nakatsuhara** is Reader in Language Assessment at the Centre for Research in English Language Learning and Assessment (CRELLA), University of Bedfordshire, UK. Her main research interests lie in the nature of co-constructed interaction in speaking tests, test development and validation, and the relationship between listening and speaking skills. She has led a number of international testing projects, working with ministries, universities, and examination boards. Her publications include the book, *The Discourse of the IELTS Speaking Test* (with P. Seedhouse, 2018, CUP), and research articles in journals such as *Language Testing*, *Language Assessment Quarterly*, *Modern Language Journal*, and *System*.

**Lyn May** is a Senior Lecturer at the Queensland University of Technology, Australia. Lyn's research interests include L2 speaking, interactional competence and learning oriented assessment. Lyn is the author of *Interaction in a Paired Speaking Test: The Rater's Perspective* and her work has been published in journals including *Language Testing*, *Language Assessment Quarterly* and *Assessment in Education: Principles, Policy & Practice*.

**Chihiro Inoue** is Senior Lecturer in Language Assessment at the Centre for Research in English Language Learning and Assessment (CRELLA), University of Bedfordshire, UK. Her expertise lies in the assessment of L2 speaking and listening, particularly in the research on test tasks, test-taking processes and salient features of learner language. She has led a number of language test development and validation projects for government bodies and testing organisations all over the world. She has published a monograph, *Task Equivalence in Speaking Tests* (Peter Lang, 2013) and in international journals such as *Language Assessment Quarterly*, *Language Learning Journal* and *Assessing Writing*.

**Edit Willcox-Ficzere** is a Subject Coordinator and Senior Lecturer at Oxford Brookes University, UK. Her main research interests lie in the teaching and assessment of L2 pragmatic competence. She has presented on these topics at a number of international conferences such as ALTE, EALTA and INPRA. Her publications include an article on the *Usefulness of Teaching Politeness Strategies in English Language Classrooms* (*Baltic Journal of English Language, Literature and Culture* 2014) and a book review on *Assessing Second Language Pragmatics* (*Language Testing* 2016).

**Carolyn Westbrook** is a Test Development Researcher at the British Council in the UK. Formerly an Associate Professor in EFL, she has over 25 years' experience teaching and assessing EAP and ESP. A Senior Fellow of the Higher Education Academy in the UK, Carolyn has been a teacher, tester, teacher trainer and materials writer. She has been interested in Language Assessment since 2007, taking part in a number of testing and assessment development projects, as well as being involved in major projects on language assessment literacy for teachers. She has also published several books and journal articles.

**Richard Spiby** works at the British Council in London with the Assessment Research Group. His responsibilities include conducting research into existing tests and developing new receptive skills components. He also has extensive experience in assessment production and management and has worked on a variety of assessment development and training projects worldwide. His areas of interest include language assessment literacy, test-taker strategies, test analysis and inclusivity in language assessment.

# CONTENTS

<b>1. RESEARCH BACKGROUND</b>	<b>5</b>
1.1 Interactional competence and pragmatic competence	5
1.2 Eliciting interactional and pragmatic competence in semi-direct tests	5
1.3 Research questions	6
<b>2. METHODOLOGY</b>	<b>7</b>
2.1 Stage 1: Task development	7
2.2 Stage 2: Piloting the tasks	8
2.2.1 Participants	8
2.2.2 Test tasks	9
2.2.3 Post-test feedback questionnaire and semi-structured interviews	10
2.2.4 Data analysis	11
<b>3. RESULTS</b>	<b>12</b>
3.1 Assessment of interactional competence (IC)	12
3.1.1 B1 performances	14
3.1.2 B2 performances	15
3.1.3 C performances	16
3.2 Assessment of pragmatic competence (PC)	17
3.2.1 B1 performances	19
3.2.2 B2 performances	19
3.2.3 C performances	20
3.3 Test-taker perceptions and feedback	21
3.3.1 IC tasks	21
3.3.2 PC tasks	22
<b>4. CONCLUSIONS</b>	<b>23</b>
<b>REFERENCES</b>	<b>24</b>
<b>Appendix A: IC and PC Task Specifications</b>	<b>26</b>
<b>Appendix B: Task instructions and prompts</b>	<b>28</b>
<b>Appendix C: Post-test survey results</b>	<b>39</b>
<b>Appendix D: Post-survey interview questions</b>	<b>46</b>
<b>Appendix E: Transcription notations</b>	<b>48</b>
<b>LIST OF TABLES AND FIGURES</b>	
<b>Table 1:</b> Average word counts in IC tasks (a) and IC(b)	12
<b>Table 2:</b> IC interactional moves: IC(a+b) Parts 1 and 2 (n = 46)	12
<b>Table 3:</b> Average word counts in PC tasks (a) (b)	17
<b>Table 4:</b> PC interactional moves: PC(a+b) (n = 46)	18
<b>Figure 1:</b> Selected screenshots of IC(a)	9
<b>Figure 2:</b> Selected screenshots of PC(b)	10
<b>Figure 3:</b> IC (a+b) Parts 1 and 2: Interactional moves across B1, B2 and C performances	14
<b>Figure 4:</b> PC (a+b): Interactional moves across B1, B2 and C performances	18

# 1. RESEARCH BACKGROUND

---

## 1.1 Interactional competence and pragmatic competence

Galaczi and Taylor (2018, p. 226) define interactional competence (IC) at the broad macrolevel as “the ability to co-construct interaction in a purposeful and meaningful way, taking into account sociocultural and pragmatic dimensions of the speech situation and event”. This understanding of IC overlapping with pragmatic competence (PC) is clarified by Young (2019), who argues:

“[while] L2 pragmatics research addresses the pragmatic meanings that speakers and hearers attribute to the conventional and context-specific meanings of utterances, the reach of IC research goes beyond pragmatics in its recognition of IC as co-constructed by all participants in a discursive practice rather than simply the knowledge or competence of a single speaker or hearer”. (p. 106)

That is, pragmatic competence is the ability that resides within a speaker to use language efficiently in a contextually appropriate fashion, whilst “IC is not the knowledge or the possession of an individual person but is co-constructed by all participants in a discursive practice, and IC varies with the practice and with the participants” (ibid., p. 96).

## 1.2 Eliciting interactional and pragmatic competence in semi-direct tests

The affordances of technological advances for the assessment of L2 speaking are well documented (e.g., Lim, 2018). There are now a wide range of computer-delivered semi-direct speaking assessments, including computer-based integrated-skills assessment (e.g., TOEFL iBT), and computer-delivered and machine-scored assessment (e.g., PTE Academic, Cambridge Linguaskill). However, despite its practical advantages, a semi-direct test, which does not permit reciprocal interaction between speakers, is limited in the range of constructs that it can tap into. In the final paper of the 2018 *Language Testing Special Issue* on the assessment of IC, Plough, Banerjee and Iwashita (2018) note that ways in which the co-constructed nature of IC can be operationalised in a semi-direct test continue to challenge the field.

Recently, the potential of Automatic Speech Recognition (ASR) and Spoken Dialogue Systems (SDSs) to generate ‘spoken interaction’ with an L2 test-taker has started attracting our attention as a possible way to expand the construct of semi-direct speaking tests in the future. For example, Litman et al.’s (2016) proof-of-concept research investigated the feasibility of using an existing SDS to interact with, and collect interactional data from, L2 speakers. Timpe-Laughlin et al. (2017) explored how an SDS that allows multi-turn conversations by generating responses based on prompts can be used to interact with L2 speakers on a pragmatic task. While these initiatives are promising, the use of the currently available SDSs with spontaneous speech by L2 speakers containing incorrect lexis and grammatical structures and reduced intelligibility still poses challenges for the system to respond naturally and appropriately (Litman, Strik & Lim, 2018).

It should also be noted it is unlikely that a SDS-based semi-direct speaking test will become accessible in the near future to those examination providers with limited resources to invest in setting up and training the system.

As such, a general consensus for the moment remains to be that “computer-delivered speaking tests are unidirectional and lack the element of co-construction”, with the performance being “typically elicited through technology-mediated prompts and the conversation has a pre-determined course which the test-taker has no influence on” (Galaczi & Taylor, 2018, p. 232). While acknowledging that monologic tasks contextualised as speaking to an audience entail recipient design of talk, Roever and Kasper (2018, p. 350) caution that they “severely under-represent the IC construct”.

However, there are aspects of IC such as providing contingent responses (Roever & Kasper, 2018), interactive listening (Lam, 2018; Ross, 2018), persuading, and asking for/providing clarification (van Batenburg, Oostdam, van Gelderen & de Jong, 2018) that may be elicited through semi-direct tests of speaking, which are worth experimenting with in a semi-direct format. These features of IC are indeed identified as salient when examiners rate interactional skills of test-takers in a face-to-face paired discussion test (May, Nakatsuhara, Lam & Galaczi, 2020).

On the other hand, the assessment of pragmatic competence (PC) seems more plausible under a semi-direct, monologic task condition, since reciprocal interaction between speakers does not seem as essential as in the assessment of IC. Notable examples include English Language Testing’s Password Speaking Test<sup>1</sup>, which has a task that requires test-takers to perform a particular speech act and demonstrate their pragmatic competence. Indeed, Ikeda’s (2017, p. 270) study of L2 pragmatics assessment using both monologic and dialogic speaking tasks reports that “test takers’ performances under the dialogue condition with the presence of an interlocutor and those under the monologue condition without interlocutor were highly correlated”. Nevertheless, given the importance of the detailed context in which the learner’s utterance is produced, the extent to which limited on-screen instructions can provide comprehensive contextual information for the test-takers to work with remains a challenge. Rich, fine-grained contextual cues are especially significant when context is conceptualised as embracing both the speaker and listener, as in Taguchi and Rover’s (2017, p. 1) definition: “[the] field of pragmatics studies aspects of language systems that are dependent on the speaker, the listener, and the context of an utterance”. The use of visuals to provide contextual cues is thus a promising avenue to pursue in the design of assessment tasks.

### 1.3 Research questions

The main objective of this research is therefore to explore widening the construct measured in a semi-direct speaking test, by developing tasks that could elicit selected features of IC and PC. In so doing, we do not use advanced technologies that resource-limited examination providers may find difficult to access. Underpinning the exploration of the new IC and PC tasks were two research questions:

- RQ1.** To what extent can a computer-based speaking test elicit features of interactional competence at different levels?
- RQ2.** To what extent can a computer-based speaking test elicit features of pragmatic competence at different levels?

---

<sup>1</sup> <https://www.englishlanguagetesting.co.uk/test-format/password-speaking-sample-questions/>

## 2. METHODOLOGY

---

The research consisted of two stages:

1. developing two tasks to elicit interactional competence and two tasks to elicit pragmatic competence
2. piloting these four tasks with 48 test-takers.

This report focuses mainly on the second stage of the project, which sought evidence of the use of interactional and pragmatic competence through the analysis of test-taker output language elicited by the four tasks and their test-taking experience and perceptions of the tasks.

### 2.1 Stage 1: Task development

We first established guiding principles in developing the IC and PC tasks, and specified the elements of IC and PC constructs which we wished to target under the computer-delivered condition. The discussions were informed by recent IC and PC literature, as well as our own empirical studies (e.g., Ficzer, 2019; Galaczi & Taylor, 2018; Ikeda, 2017; May et al., 2020; Taguchi & Roever, 2017; Youn, 2015).

The main guiding principles were as follows:

- a. In order to enhance the authenticity of the tasks and to provide a rich context, the use of visuals and videos is essential (Ginther, 2002).
- b. For the IC tasks, a generic clarification request is included after the first part of the test-taker response. An effective response to this request could provide evidence of interactive listening skills. The two IC tasks focus on giving and justifying opinions, acknowledging and disagreeing with a (hypothetical) partner's opinion and persuading them to change their opinion.
- c. The two PC tasks focus on the speech act of 'request'. The reason for this is twofold. On the one hand, the frequent occurrence of requests in real life makes them not only important but also familiar to L2 speakers (e.g., Ficzer, 2019). In addition, their inherently longer sequences allow a deeper insight into speakers' use of pre-expansion moves that lead to the request (e.g., Al-Gahtani & Roever, 2012; Lee, 2009). The level of imposition in both tasks is relatively high, as these contexts tend to elicit more pre-expansion moves. The power relationship between the speaker and listener is equal in one task, and unequal in the other so as to test how and whether the speakers vary their pre-expansion moves in the different contexts (e.g., Ficzer, 2019; Ikeda, 2017; Taleghani-Nikazm & Huth, 2010; Youn, 2015).
- d. All input language in task instructions and prompts is standardised. The target output level of the tasks is set to B2 of the CEFR (Council of Europe, 2001), and all audio and written input language is set to B1 or below (n.b., all exceptions are carefully discussed). For that purpose, English Vocabulary Profile and English Grammar Profile are used ([www.englishprofile.org/](http://www.englishprofile.org/)). The speed of audio input is also controlled.
- e. In order not to hinder the demonstration of test-takers' IC and PC skills due to a lack of understanding of listening input and/or working memory capacity, written texts are displayed on screen for scaffolding purposes. We trial tasks prior to the pilot test to determine what to present on screen.
- f. We use the auto-play mode of Microsoft PowerPoint for delivering tasks, so that the participants do not need to install a special program to run the test.



Following these principles, the task specifications were developed using Weir's (2005; further elaborated in Taylor, 2011) socio-cognitive framework. As detailed in the task specifications in Appendix A, the IC tasks measure test-takers' ability to:

- disagree and put forward a different point of view effectively and with justification
- acknowledge partner's views
- effectively link their own contribution to the partner's contribution
- work towards a decision by trying to persuade the partner
- demonstrate they have been listening carefully/attentively through responding appropriately to the partner's idea and being able to clarify/rephrase their points.

The PC tasks were designed to measure test-takers' pragmatic competence in making requests, focusing on the test-takers' ability to:

- structure a coherent sequence, including logical pre-expansion of features such as projecting the upcoming request and providing a reason for the request, which achieves the communicative goal of 'requesting'
- connect the different pre-expansion features appropriately (e.g., using conjunctions) in order to aid the listener's comprehension of the entire communicative act
- deliver the intended meanings while being sensitive to the social and power relationship between interlocutors, imposition and showing awareness of these contextual factors in their language choices.

## 2.2 Stage 2: Piloting the tasks

### 2.2.1 Participants

For the pilot test, a total of 48 test-takers were recruited from China (ID: C01-C24) and Austria (ID: A25-48). They performed all four tasks as well as one benchmarking task. As explained in the next section, their proficiency levels were determined by the benchmark task performances: 15 test-takers at B1, 15 at B2 and 16 at C (n.b., since the benchmarking task we used is taken from the British Council's Aptis Speaking Test which does not differentiate C1 and C2, this study used an overall C category.). As the ratings of two test-takers were not fully agreed among raters, the spoken data of these two test-takers were excluded from our output language analysis, while their survey data were included. The 48 participants consisted of 27 males and 21 females, who were either high-school or university students, with their ages ranging from 18 to 30 years (Mean: 20.5, SD: 2.8). Approximately half of the test-takers had a European first language (L1) background and the rest had an Asian L1 background (i.e., 22 German, 1 Croatian, 1 Turkish, and 24 Mandarin Chinese L1 speakers).



## 2.2.2 Test tasks

The test included five tasks: one benchmarking, two IC tasks, and two PC tasks. All participants took the benchmarking task as the first task, followed by the remaining four tasks. The order of the IC and PC tasks was counter-balanced. All task performances were audio-recorded.

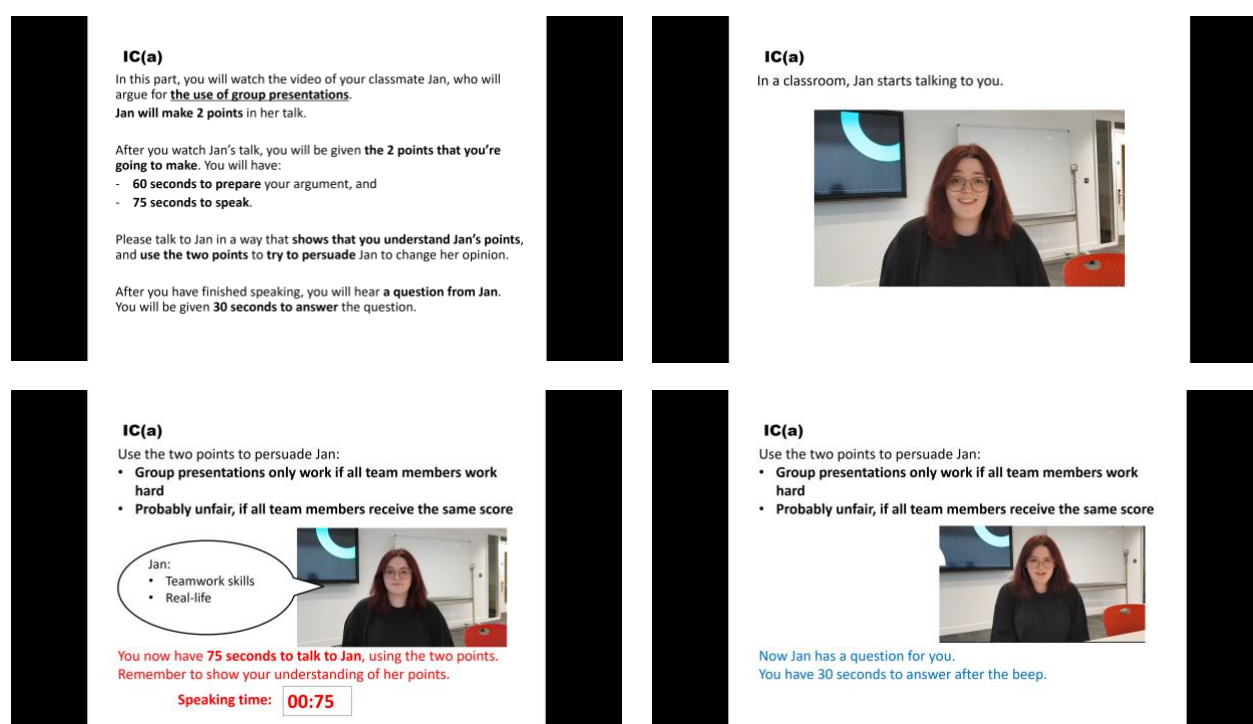
### Benchmarking task

A Part 4 task from the Aptis Speaking Test (O'Sullivan et al., 2020) was used as a benchmarking task. In this task, after one-minute preparation time, test-takers gave a long turn of two minutes, integrating responses to a set of three questions.

### Two IC tasks – IC(a) and IC(b)

In both tasks, test-takers watched a video in which a conversation partner expressed views in 140 to 160 words. They were then asked to express a differing opinion while appropriately responding to and persuading the partner, and also using interactive listening skills. Key words from the partner's talk, as well as the points that the test-takers were required to make, appeared on screen. One-minute preparation time was given and the response time lasted 75 seconds. After the completion of the response time, an automated video was played to ask a standardised clarification question (i.e., *Sorry, I didn't get the last point. Can you say that again, perhaps using different words or giving an example?*), to which the test-takers were required to respond in 30 seconds, without preparation time. The topics of IC(a) and IC(b) were the pros and cons of participating in group versus individual presentations, and of travelling alone versus with friends, respectively. Selected screenshots of IC(a) are presented in Figure 1 (see Appendix B for the full task instructions and prompts of IC(a) and IC(b)).

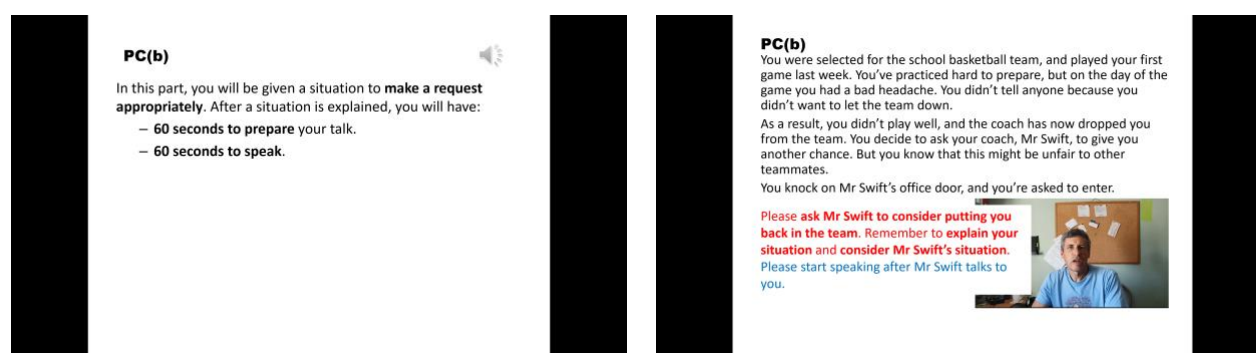
Figure 1: Selected screenshots of IC(a)



## Two PC tasks – PC(a) and PC(b)

The test-takers were given a situation to recognise the pragmatic demands underlying the task. All input was aural, with textual support provided on screen. They had one-minute preparation time, followed by a short video of a conversational partner asking a question (i.e., *Hi, is everything OK? You look a bit worried* in PC(a); *Hello. How can I help you?* in PC(b)). Test-takers then had one minute to initiate, develop and close a sequence to achieve the communicative goal of 'request'. PC(a) was to request the listener to perform a wish (a classmate to exchange a presentation time with the speaker) and PC(b) was to request the listener to grant a wish (a basketball team coach to consider putting the speaker back in the team). Selected screenshots of PC(b) are presented in Figure 2 (see Appendix B for the full task instructions and prompts of PC(a) and PC(b)).

**Figure 2:** Selected screenshots of PC(b)



## 2.2.3 Post-test feedback questionnaire and semi-structured interviews

Immediately after completing the test, all test-takers completed a feedback questionnaire, and 12 test-takers were further interviewed by one of the trained test administrators. The questionnaire and interviews were given in German for the Austrian cohort and in Mandarin Chinese for the Chinese cohort. As shown in Appendix C (English version), the questionnaire consisted of:

- 6 demographic questions
- 9 selected response items about the IC tasks (e.g., length of preparation and response time, clarity of intended audience, task authenticity, usefulness of video input)
- 11 selected response items about the PC tasks (e.g., length of preparation and response time, clarity of intended audience, task authenticity, power relationship between the speaker and listener, level of imposition, ease of making the request)
- 2 free comment boxes, one for the IC and one for the PC tasks.

Semi-structured interviews were carried out with six participants from the Chinese cohort and six participants from the Austrian cohort, consisting of five B1, three B2 and four C test-takers. Upon completion of the questionnaire, trained test administrators quickly skimmed their questionnaire responses, and interviewed them to gain further insights into their questionnaire data. They had 12 pre-scripted interview questions linked to the questionnaire items (see Appendix D), from which relevant questions were selected according to the test-takers' responses. Each interview took approximately 20 minutes, and all interviews were audio-recorded.

## 2.2.4 Data analysis

### Rating benchmarking task performances

Forty-eight audio-recorded performances on the benchmarking task were rated by two experienced Aptis Speaking raters using the Aptis Part 4 rating scale that covers A1/A2 (Score 0), B1 (Scores 1 and 2), B2 (Scores 3 and 4), and C (Scores 5 and 6) (O'Sullivan et al., 2020). When the two raters did not agree, a senior rater trainer served as a third rater. Since there was still a discrepancy in rater agreement for two performances even after the third rater was involved, we decided to exclude those two test-takers' spoken data from our analysis.

It should be noted that we did not intend to advocate that learners' development of IC and PC is aligned with other language aspects measured by the long-turn benchmarking task. However, it was believed that the broad categorisation of the three CEFR levels (B1, B2, C) based on the ratings of the benchmarking task would still be useful when exploring criterial features of the three broad levels.

### Interactional move analysis

The audio-recorded responses to the IC and PC tasks were transcribed (see Appendix E for the transcription notations), and a total of 184 transcripts (i.e., 46 test-takers x 4 tasks) were coded for moves for different communicative purposes (e.g., Ficzere, 2019; Young, 2011). The identification of moves was conducted both deductively and inductively, involving repeated coding sessions and inter-coder reliability checks.

1. Consulting a provisional list for IC and PC moves drawing upon May et al. (2020) and Ficzere (2019), one of the researchers coded six IC transcripts and another researcher coded six PC transcripts, based on which initial IC and PC coding schemes were developed.
2. The whole project team discussed the initial coding schemes while coding three IC and three PC transcripts together, and made several changes to the initial coding schemes.
3. Two researchers then coded four IC transcripts individually, and coder-reliability was calculated. All discrepancies were discussed until agreement was reached, and further modifications were made to the coding schemes. The same procedure was applied to the PC coding by two other researchers.
4. The two groups of two researchers then coded another set of four IC and four PC transcripts, respectively, and a second round of coder-reliability was calculated. Since the coder agreement was still insufficient for PC (75.6%), discrepancies were discussed, and further adjustments were made to the PC coding scheme. This was followed by a third round of calculating PC coder-reliability with a new set of four PC transcripts.

The agreement rates obtained in the second coder-reliability check for IC and the third check for PC were 99.2% and 95.4%, respectively. Upon the achievement of sufficiently high agreement rates, the remaining PC and IC transcripts were single-coded by the two researchers (one focusing on IC and the other on PC) who led the development of each coding scheme and who were also involved in the final coder-reliability checks. The final IC and PC coding schemes are presented in the Results section. The coding outcomes were then quantitatively and qualitatively compared across three proficiency groups.

### Post-test questionnaire and interview analysis

All responses to selected response items were analysed quantitatively. While the small sample size did not allow any inferential statistics, descriptive statistics were run for the entire data set of 48 test-takers, and responses by two cohort groups (i.e., one in Austria and the other in China) and by three proficiency groups (i.e., B1, B2 and C) were also compared in case there were notable between-group differences.

The 12 audio-recorded interviews were transcribed and translated into English, and the transcripts and all comments provided in free comment boxes in the questionnaires were thematically analysed by two researchers. Since all interview questions were linked to questionnaire items, it was straightforward to identify themes in the interview data. First, one researcher identified 14 themes and coded all data accordingly. The other researcher then went through all coded transcripts and confirmed accuracy. The two researchers then discussed the coding together to make only a few minor adjustments. As the dataset was small and coding straightforward, no special data management software was necessary.

## 3. RESULTS

### 3.1 Assessment of interactional competence (IC)

Table 1 presents the number of words the two IC tasks elicited on average from the three levels of test-takers, showing the steady increase in word counts both in the initial response part (Part 1) and in the clarification part (Part 2) of the tasks. It is noteworthy that while producing the least average number of words among three proficiency groups, B1 test-takers were able to produce substantial responses both in Part 1 (average 114 words, response time 75 seconds) and Part 2 (average 47 words, response time 30 seconds) of the tasks.

**Table 1:** Average word counts in IC tasks (a) and IC(b)

Level	IC task	Part 1 Mean (SD)	Part 2 Mean (SD)	Total Mean (SD)
<b>B1</b> (n=15)	(a)	113.53 (39.50)	47.00 (19.75)	161.73 (53.99)
	(b)	115.73 (36.17)	47.20 (20.89)	
<b>B2</b> (n=15)	(a)	144.07 (25.80)	54.73 (17.87)	200.30 (36.94)
	(b)	145.13 (29.69)	56.67 (9.56)	
<b>C</b> (n=16)	(a)	162.50 (20.63)	67.69 (19.65)	236.09 (33.06)
	(b)	174.25 (25.02)	67.75 (12.27)	

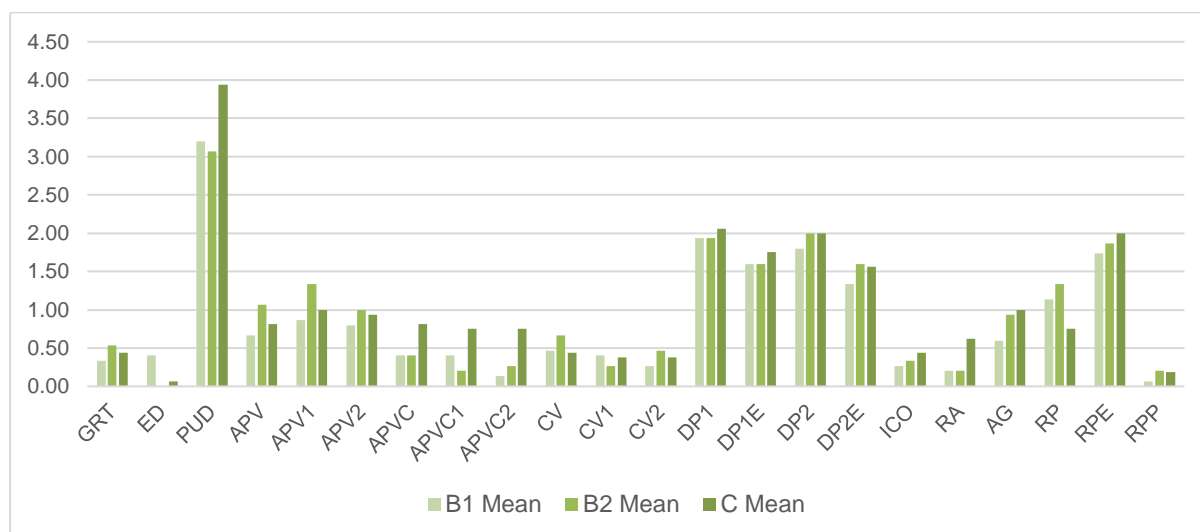
The analysis of test-taker talk revealed that a range of interactional moves were elicited at all levels. Table 2 shows a full list of coded IC interactional moves with the total number of segments identified for each code as well as means (i.e., average use per person) and SDs across the three proficiency levels. The interactional moves were identified deductively through those included in the task instructions (e.g., *acknowledging the speaker's point of view*, *disagreeing using the first point from the task prompt*) and inductively through close analysis of the transcripts of test-taker responses (e.g., *greeting*, *projecting upcoming disagreement*).

**Table 2:** IC interactional moves: IC(a+b) Parts 1 and 2 (n = 46)

MOVES USED TO RESPOND TO THE INITIAL QUESTION (Part 1)	ABBR	TOT	B1 M(SD)	B2 M(SD)	C M(SD)
Greeting	GRT	20	0.33(0.62)	0.53(0.74)	0.44(0.73)
Explicit disagreement	ED	7	0.40(0.74)	0.00(0.00)	0.06(0.25)
Projecting upcoming disagreement	PUD	157	3.20(1.47)	3.07(0.96)	3.94(1.00)
Acknowledging speaker's overall point of view/opinion/experience on the topic	APV	39	0.67(0.82)	1.07(0.59)	0.81(0.83)
Acknowledges/restates speaker's point of view in relation to the first or second point made by the speaker	APV1	49	0.87(0.74)	1.33(0.72)	1.00(0.82)
	APV2	42	0.80(0.77)	1.00(0.93)	0.94(0.77)
Appears to acknowledge and concur with speaker's point of view in general, but this is preparing the ground to be retaken in a subsequent counter-move	APVC	25	0.40(0.63)	0.40(0.51)	0.81(0.83)
Appears to acknowledge and concur with speaker's first or second point, but this is preparing the ground to be retaken in a subsequent counter-move	APVC1	21	0.40(0.63)	0.20(0.41)	0.75(0.68)
	APVC2	18	0.13(0.35)	0.27(0.70)	0.75(0.77)
Counter-view/perspective on general argument	CV	24	0.47(0.64)	0.67(0.72)	0.44(0.73)
Counter-view/perspective on either point 1 or 2 made by the speaker	CV1	16	0.40(0.63)	0.27(0.59)	0.38(0.50)
	CV2	17	0.27(0.46)	0.47(0.83)	0.38(0.72)
Disagrees, using first point from the task prompt	DP1	91	1.93(0.59)	1.93(0.26)	2.06(0.25)
Elaborates on the first point of disagreement	DP1E	76	1.60(0.74)	1.60(0.63)	1.75(0.68)
Disagrees, using second point from the task prompt	DP2	89	1.80(0.56)	2.00(0.00)	2.00(0.00)
Elaborates on the second point of disagreement	DP2E	69	1.33(0.90)	1.60(0.63)	1.56(0.81)
Invitation/suggestion to change opinion/agree with the overall argument presented	ICO	16	0.27(0.46)	0.33(0.62)	0.44(0.63)
Reiterates/summarises overall argument	RA	16	0.20(0.41)	0.20(0.56)	0.63(0.72)
<b>MOVES USED TO RESPOND TO A CLARIFICATION REQUEST (Part 2)</b>					
Agrees to clarify/restate last point	AG	39	0.60(0.83)	0.93(0.96)	1.00(0.73)
Restates last point	RP	49	1.13(0.74)	1.33(0.62)	0.75(0.68)
Exemplifies/elaborates on the last point	RPE	86	1.73(0.59)	1.87(0.35)	2.00(0.00)
Invites/persuades interlocutor to change/question their view on last point	RPP	7	0.07(0.26)	0.20(0.56)	0.19(0.40)

Figure 3 visually represents the usage of each move across the three proficiency levels. It is important to note that there were a range of performances within each level, in terms of the IC features noted. However, certain IC features were more apparent in a level, which the following sections will explore and exemplify. For example, the use of *projecting upcoming disagreement* (PUD) by test-takers at C level (mean: 3.94) was greater than the B1 and B2 groups (mean: 3.20 and 3.07, respectively). Similarly, utterances that *appear to acknowledge and concur with speaker's point of view* (APVC) were more frequently observed in the responses of C-level test-takers than those of B-level test-takers (Means: B1=0.40, B2=0.40, C=0.81). APVC1 and APVC2 also followed the same pattern.

**Figure 3:** IC (a+b) Parts 1 and 2: Interactional moves across B1, B2 and C performances



### 3.1.1 B1 performances

Students at B1 level were generally able to structure their speech, project upcoming disagreement and sometimes acknowledge the hypothetical interlocutor's point of view, but not as often as this occurred in B2 and C performances. Concurring concessions, where the test-taker appears to acknowledge and concur with the hypothetical interlocutor's point of view in order to prepare the ground to be retaken in a subsequent move, were used but less often than a direct acknowledgement. The initial statement of their point of view was often expressed abruptly through an explicit statement of disagreement. B1 test-takers were generally able to state a differing opinion and provide limited support for this opinion, using the points and often exact wording from the prompt. They were also able to express counter views. In Excerpt 1 the test-taker begins by expressing explicit disagreement with Dan's point of view (line 2) and projects upcoming reasons (line 3) before stating the differing opinion and providing support (lines 4 and 5). This test-taker does not acknowledge Dan's point of view.

#### Excerpt 1: Task IC(b) Part 1 / B1 sample (ID: C06)

- 1 (erm) In my opinion, (er) (.) in my opinion,
- 2 → I disagreed with you,
- 3 → (er) because I think
- 4 → (er) go, (.) go travelling with my, with my friends is better,
- 5 → because (er) you can talk, (.) you can talk with your friends during a journey, (er) if you, if
- 6 you are, (..) if you feel boring.

In response to the request to clarify or provide an example of the last point, B1 test-takers were less likely to explicitly agree to the request. This is important to note, as responding by agreeing could be interpreted as demonstrating interactive listening. The typical response at this level was to repeat the last point, sometimes verbatim, with some test-takers also providing an example, but not as often as B2 and C test-takers. In Excerpt 2, the test-taker repeats the last point (lines 1–2) and then attempts to give an example (lines 3–4), which lacks clarity, possibly because of a limited vocabulary range.



### Excerpt 2: IC(b) Part 2 / B1 sample (ID: C17)

- 1 My last point (er) (..) (er) (..) you, you, you can, (..) you can be cheaper to share costs with
- 2 (er) friends, (uh) (..) (er) which I said,
- 3 → (uh) (..) (er) like (er) (..) eating cost. (uh) (..) You can (..) try to ( ) (..) for each other, (..)
- 4 which is (er) be cheaper than (..) your (..) travel alone (..) to eat.

A focus on the listener was evident in some B1 performances; however, this was much more pronounced in performances from higher level learners. While greeting and acknowledging the hypothetical interlocutor's point of view were clear aspects of orienting to them as listener, this was also done through countering their view and the use of the pronoun 'you' to refer to the hypothetical interlocutor. Examples include *So, I r-, recommend (er) you to (..) (er) presentation by yourself* (C08) and *you mentioned that you can change plans* (A45).

### 3.1.2 B2 performances

Performances from students at B2 level were more likely to greet the hypothetical interlocutor and tended to include acknowledgement of their point of view, sometimes expressed as a concurring concession. Points were elaborated, contained more sophisticated lexis, and there were more counterviews of the overall points than in B1 performances, which made responses at this level more persuasive.

In the excerpt below, test-taker A46 begins by greeting Jan, and follows this by acknowledging Jan's views through a series of concurring concessions in lines 2, 3 and 4. This test-taker projects the upcoming disagreement in line 6 through the effective use of 'in the worst case scenario', followed by the statement of opinion and elaboration.

### Excerpt 3: IC(a) Part 1 / B2 sample (ID: A46)

- 1 Hey, Jan.
- 2 → (erm) (..) That's a great opinion you have
- 3 → (..) I also share this opinion with you, (..) to work in a team, (..) and (..) get teamwork skills
- 4 → (..) and (..) (erm) (..) develop (..) you (..) skills (than) (..) (er) for the real life (..) in group
- 5 presentation.
- 6 → (..) but in the worst case (..) scenario
- 7 (..) (er) you get (..) a (..) teammate (..) who (..) will not work (..) at all (..) or isn't interested (..) in the
- 8 project (..) and you have to do it alone
- 9 (..) and d-, d-, (er) there is a risk (..) that that happened, (..) can happen.

In response to the request to clarify and/or elaborate on their last point in Part 2 of the task, B2 test-takers were more likely to explicitly agree to the request and provide both a restatement of their last point and an example. This can be seen in the excerpt below where C19 responds to the request in line 1, provides an extended example in lines 2–5 and restates the final point in line 6.

### Excerpt 4: IC(b) Part 2 / B2 sample (ID: C19)

- 1 → (er) (..) Yeah.
- 2 → (erm) (er) Let's have an example. (er) Maybe you will choose a, (..) choose a (..) single room with one
- 3 hundred dollars, (..) one hundred dollars. But, however, if you (er) travel with your friends, you can
- 4 just pay, (erm) you can just, (er) (er) you two can just pay one hundred and fifty dollars, (..) (erm) in
- 5 a room.
- 6 → So, (er) it will make you (..) pay much cheaper than before.



A focus on the listener became more evident in B2 performances. A test-taker may explicitly address the hypothetical interlocutor during Part 1, with C05 following up with Jan in IC(a) on a point she has made regarding teamwork and asking her to reconsider: *However, I want to say that, (er) (...) Jan do you think it's, (...) it's sometimes teamwork, only (...) (er) work if all team members work really hard?* and C10 asking Dan in IC(b): *But have you ever thought about travelling together (...) also has (...) many advantages?* The hypothetical interlocutor may be asked to change their mind in a way that clearly orients to them as listener, as we can see in C19's wrapping up of Part 1: *So, I believe you, (...) if you travel with friends, (er) you will be happier and, (...) (erm) you will be happier and (...) get more, (erm) get more unforgettable experience.*

### 3.1.3 C performances

Test-takers at a higher level were more likely to acknowledge and then refute the hypothetical interlocutor's point of view through a concurring concession. They effectively countered views, in addition to putting forth the points of disagreement, with the structure of their arguments often being more complex and persuasive than those of test-takers at lower levels. Shades of meaning were communicated through statements such as *I agree with you to some extent* (C16) and upcoming disagreements were clearly flagged more often than at other levels. The other feature more apparent in C performances was test-takers summarising/reiterating the overall argument at the end of Part 1: *And so, I would (...) rather (erm) (...) recommend this* (A27).

In Excerpt 5, after greeting Dan, A36 acknowledges Dan's second point of view with an extended concurring concession (lines 2–5), acknowledges the first point of view and then projects upcoming disagreement in a way that clearly orients to persuading Dan (line 15). A counterview is clearly established through a comparison in lines 11–12, followed by another point of disagreement and then an invitation for Dan to change his mind (line 15).

#### Excerpt 5: IC(b) Part 1 / C sample (ID: A36)

- 1        Yeah, Dan.
- 2        → I remember, (erm) you told me about your Scotland trip. And I can totally see your point that (...)
- 3        (erm) if you travel on your own, you (...) will be (...) much more likely to (...) meet local people,
- 4        because you, yourself have to get (...) (erm) active, have to become active (...) and to (...) introduce
- 5        yourself to others, to (erm) (...) I don't know, talk to them in a bar or something.
- 6        (er) (...) and also that you're more flexible with your (erm) (...) trip choices or with the time you
- 7        want to get up. (er)
- 8        → While all of that is true, and is super important, (erm) you must not forget
- 9        that (er) with friends, it can be a lot of fun. So, (...) (erm) the experiences and memories you create
- 10       during this trip (...) can stay with you for a whole lifetime.
- 11       → And (...) I think that's valuable in itself and (...) maybe (erm) (...) it's more important than (...) being
- 12       super independent, (erm) (with) all your (...) choices.
- 13       And (...) also, (er) (...) if you have you have friends with you, (erm) like (...) travelling by taxi, or by
- 14       train or (er) (...) eating out, maybe it's (...) cheaper, or, (...) because you can share the costs.
- 15       → So, you should consider that as well. Maybe you want to change your mind.

In response to the request to clarify or elaborate on their last point, test-takers at C level were more likely to explicitly agree to the request and provide an example, rather than simply repeating the point. In the example below, A32 responds directly to the request in a way that is meant to make the listener feel at ease (line 1), then provides an example that clearly supports the last point (lines 2–4), and finishes by reorienting to persuading the listener (lines 5–6).

### Excerpt 6: IC(b) Part 2 / C sample (ID: A32)

- 1 → Sure. No problem.  
 2 (erm) An example would be, like renting a car. I mean if, if you (.) were to rent a car, it  
 3 could, (.) you know, cost you up to five hun-, five hundred pounds or euros. And, on the  
 4 other hand, if you shared it with friends, you know, you can just (.) save a lot of money.  
 5 → So, you should definitely consider (.) (erm) this aspect of (..) (expentures) as well, (.) when  
 6 travelling alone (.) or with friends.

A focus on the listener was pronounced in C level responses, with test-takers clearly orienting to the listener through the acknowledgement of and seeming alignment with their point of view through pronoun use reflecting a focus on 'you' and addressing them directly at the beginning of the response: *So, Dan, I've been in a very similar situation (..) as you are now* (A33); *Yeah, this is a really difficult situation you have there, Dan* (A31). This inclusion of the listener continued through responses, with Dan being asked directly to accept the test-taker's view by A31: *I think it won't be a problem to change your plans spontaneously with them s- and (er) (..) together is always bet- better (..) W- so why not?* and asked to consider other perspectives by A32: *And (...) also, have you thought about (..) you know the actual costs?*

## 3.2 Assessment of pragmatic competence (PC)

Table 3 shows the average numbers of words produced by the three levels of test-takers in two pragmatic tasks. Like the IC tasks, the amount of speech in a response time of 60 seconds becomes greater with increasing proficiency, although the increase is less sharp than that in the IC tasks.

**Table 3:** Average word counts in PC tasks (a) (b)

Level	PC(a) Mean (SD)	PC(b) Mean (SD)
<b>B1</b> (n=15)	122.60 (29.49)	138.13 (24.74)
<b>B2</b> (n=15)	133.67 (23.77)	147.27 (28.13)
<b>C</b> (n=16)	153.06 (18.17)	159.81 (18.84)

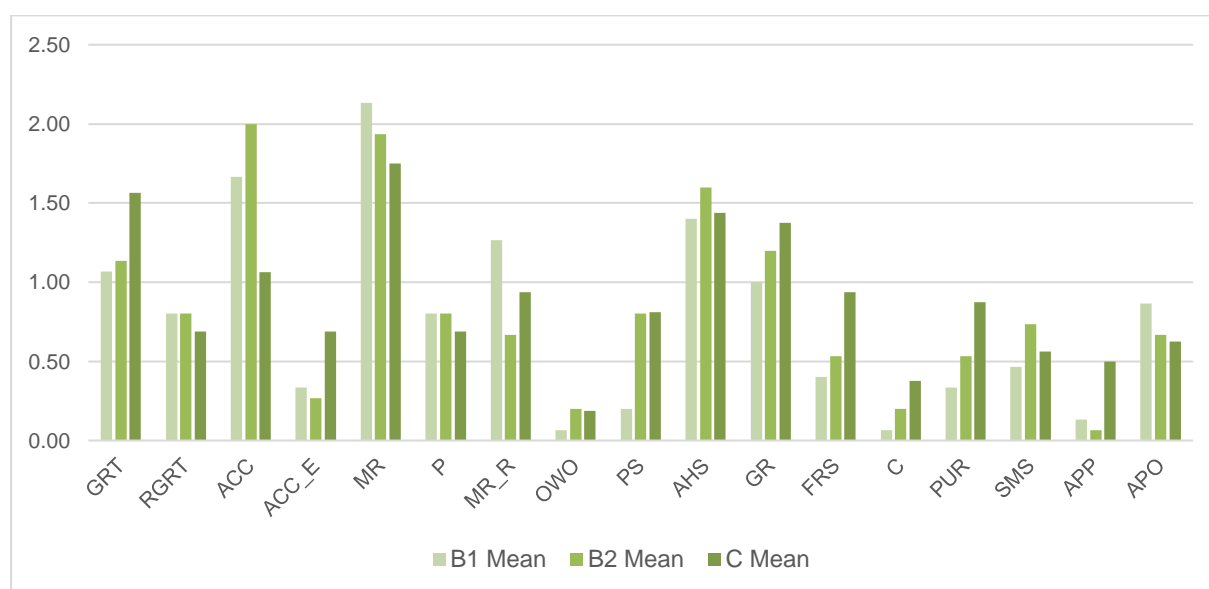
In Table 4, a list of PC interactional moves is presented with the total number of segments identified for each code as well as means and SDs across the three proficiency levels. Figure 4 visually compares the usage of PC interactional move across B1, B2 and C levels.

Table 4 and Figure 4 indicate that all test-takers employed various moves when sequencing their speech, albeit to differing degrees. For instance, *projecting the upcoming request* (PUR), involving expressions such as *Can I talk to you for a minute?* which could often preface requests in real-life interactions, was noticeably more common at C level (mean: 0.88) than at either B2 (mean: 0.53) or B1 (mean: 0.33). Similarly, *problem statements* (PS), whose role is to state the problem in order to clarify why the request has to be made, were much more commonly used at B2 and C (mean: 0.80 and 0.81 respectively) than at B1 (mean: 0.20).

**Table 4:** PC interactional moves: PC(a+b) (n = 46)

	ABBR	TOT	B1 M(SD)	B2 M(SD)	C M(SD)
Greeting	GRT	58	1.07(0.70)	1.13(0.64)	1.56(0.73)
Responding to greeting	RGRT	35	0.80(0.56)	0.80(0.56)	0.69(0.60)
Giving an account of the situation	ACC	72	1.67(0.72)	2.00(0.76)	1.06(0.77)
Giving an account with elaboration	ACC_E	20	0.33(0.49)	0.27(0.46)	0.69(0.60)
Main request	MR	89	2.13(0.35)	1.93(0.59)	1.75(0.45)
Trying to persuade listener to perform R	P	35	0.80(0.77)	0.80(0.86)	0.69(0.87)
Repeating the main request	MR_R	44	1.27(1.16)	0.67(0.62)	0.94(1.18)
Offering a way out for the listener not to perform the request	OWO	7	0.07(0.26)	0.20(0.41)	0.19(0.40)
Problem statement	PS	28	0.20(0.41)	0.80(0.77)	0.81(0.66)
Acknowledging the listener's situation	AHS	68	1.40(0.91)	1.60(0.51)	1.44(0.73)
Giving a reason for the request (referring to future event/consequence)	GR	55	1.00(0.85)	1.20(1.01)	1.38(0.62)
'Face' related statement	FRS	29	0.40(0.51)	0.53(0.74)	0.94(1.06)
Closing the message	C	10	0.07(0.26)	0.20(0.56)	0.38(0.62)
Projecting upcoming request	PUR	27	0.33(0.49)	0.53(0.52)	0.88(0.72)
Seeking mutual solution	SMS	27	0.47(0.52)	0.73(0.59)	0.56(0.51)
Expressing appreciation	APP	11	0.13(0.35)	0.07(0.26)	0.50(0.82)
Expressing apology	APO	33	0.87(0.92)	0.67(0.90)	0.63(0.81)

**Figure 4:** PC (a+b): Interactional moves across B1, B2 and C performances



Overall, there was an increasing variety of interactional moves produced as proficiency increased. B1 participants mostly included basic features that were also stated in the task instructions (e.g., *providing an account*), whereas B2 and especially C participants widely included additional features not mentioned in the instructions (e.g., PUR). Moreover, C participants often employed *face-related statements* seemingly to pre-empt the listener's negative response to the request. The following extracts exemplify how speech was typically structured at each level and highlight how test-takers' discourse was increasingly more naturally structured as proficiency increased.

### 3.2.1 B1 performances

Although all B1 test-takers attempted to structure their speech, the main request was often produced abruptly with few preliminary moves before the request. For instance, in Excerpt 7, the test-taker opens in line 1 with a somewhat awkward response to the initial greeting: *Hi, is everything OK? You look a bit worried.* This is followed by an account (lines 2–3) containing some repetition, perhaps an attempt to self-correct. The main request is then produced (lines 4–5) followed by persuasion (line 6) and another attempt to express the request (line 7), albeit incompletely. Another noticeable feature is the use of frequent fillers, hesitation and pauses despite the 60-second planning time, which may be indicating that having to formulate language as well as considering the impact of this language on the listener may have added an extra layer of difficulty.

#### Excerpt 7: PC(a) / B1 sample (ID: C24)

- 1        Yes, Connie,
- 2        → I feel very bad. (..) (erm) you have to run to the, (..) do-, (er) see the doctor, because my (.)
- 3        (er) si-, (er) I'm sick, I was sick. (..)
- 4        → So, (..) (er) could you (.) n-, ex-, (..) exchange with me for the tomorrow's (..) (er)
- 5        presentation (.) at (er) (.) 3PM?
- 6        You know, (er) (.) its won't (erm) affect your (scores).
- 7        → Just to (..) exchange (..) (er) (erm) (..),

Similar features can be observed in Excerpt 8. The test-taker starts rather abruptly with the request (line 2) without any preliminary moves and continues with verbalising the reason (line 3). However, this is left incomplete and the listener's situation is acknowledged instead (line 4). As in Excerpt 7, a number of fillers and pauses (e.g., lines 3–4) are noticeable in speech.

#### Excerpt 8: PC(b) / B1 sample (ID: C12)

- 1        Hello
- 2        → I want to, ((clears throat)) I want to (.) (erm) back in the team
- 3        → because (er) (..) I, (erm) my, (..)
- 4        (erm) **maybe** you will feel angry about that, because my (..) (er) faul-, (.) fault. (er)

### 3.2.2 B2 performances

B2 test-takers tended to sequence speech somewhat more naturally. For instance, in Excerpt 9 the test-taker first responds briefly to the greeting (line 1). What follows in line 2 seems to complement this by referring to an issue but without providing specific detail. The subsequent apology in line 3 precedes the statement projecting the upcoming request (line 4) and a detailed account (lines 5–6), after which the request is uttered in lines 7–8. The sequencing of these features seems to be less random; moves are linked more naturally (e.g., lines 4, 7–8) and the number of moves preceding the request is more substantial than at B1. Phrases such as *You know* (line 5) seem to be used naturally not only for pragmatic purposes but also for linking ideas. It was also noted that repetition was not as prevalent as in B1 speech and when it was employed it seemed to have a specific purpose, such as intensifying the message.

### Excerpt 9: PC(a) / B2 sample (ID: A47)

- 1     → No,
- 2       not everything is OK.
- 3     → I'm (..) really sorry,
- 4     → but I have to ask you something about the prepar-, (er) about the presentation, tomorrow.
- 5     → You know, I have the presentation at 3PM, and (.) you, (.) you have, (.) you present before me and
- 6       (..)
- 7     → I have to, (..) I have to ask you if it's ok if we can exchange the (..) presen- (.) -tation, so you can
- 8       start at (.) 3PM

## 3.2.3 C performances

In accordance with other studies (e.g., Youn, 2015), it was found that C level test-takers tended to supply an increasing number of supportive moves before the request. For instance, in Excerpt 10 the test-taker responds to the greeting quite naturally in line 1 and continues with a face-related statement in lines 2–3. The purpose of this statement may be twofold. It might demonstrate the test-taker's awareness of a potential obstacle (i.e., the interlocutor has a job) to the forthcoming request and it may also serve the purpose of boosting the listener's positive face in order to enhance the possibility of agreeing to perform the speaker's wish. What follows, the projection of the upcoming request (line 4), may be proof of this. Another face-related statement is added in line 5, this time seemingly to save the speaker's face, before the account is given (lines 6–7). After such substantial build up, the request is verbalised in line 8. It is also worth noting how the test-taker seems to appeal to the interlocutor by using *listen* (line 4) and *look* (line 6). The data generally indicates an increasing awareness by the test-takers of the listener's potential reactions or feelings, thus a shifting focus from 'I' to 'you' or 'we'.

### Excerpt 10: PC(a) / C sample (ID: C11)

- 1       (erm) Yes. Yes. Everything's fine with me. (erm) (..)
- 2     → I heard that you have taken a part-time job recently, congratulations. That's very independent of
- 3       you. I'm so ad-, (.) I'm so admire.
- 4     (er) Listen, I have (.) a favour to ask.
- 5     → It's kind of embarrassing, actual-, actually. (erm)
- 6       Look, we have our presentation, tomorrow, this week. (erm) (..) And (erm) (..) my origin-, (.) my
- 7       original time is 3PM, tomorrow.
- 8     → And I was (.) thinking if you can exchange that time with me,

Similarly in Excerpt 11 the test-taker does not express the request until line 9, only after having projected the upcoming request (line 2) and provided an extensive account (lines 3–7). The amount of pre-expansion is as extensive as in PC(a), which included an equal power relationship. This may be an indication that for the test-taker the level of imposition, which was equally high in this task, may be a more decisive factor when considering the sensitivity of the context than the power relationship. It is also worth noting here instances of the test-taker accepting responsibility (e.g., *It's not an apology for my worst play*), thereby undermining their own face but potentially evoking empathy in the listener, which in turn could contribute to the interlocutor granting the wish. Speech is linked naturally by the use of linking words such as 'so', 'first' and 'secondly' (lines 8–9).

### Excerpt 11: PC(b) / C sample (ID: A32)

- 1 Hello, Mr Swift.  
2 (erm) I wondered if you had a minute (.) for me to talk about the last (.) (er) week's game.  
3 (erm) I just wanted to tell you that I've (.) practised very hard and (.) on the day, we had this game, I  
4 wasn't in (..) the best con- (.) dition. And (..) I have to admit I had a very bad (.)  
5 → headache and it's (.) not an apology (.) for my, (.) for my worst play, but (..) I should have,  
6 → I should have told you prior. (.) The ultimate thing is I didn't want to (.) let our team down.  
7 So, this is why I (.) tried to (.) give my best and (..) eventually I didn't succeed.  
8 So, (.) now, (.) first I wanted to apologise  
9 → and secondly, I (.) really wanted to ask you if you, (.) if you could (.) consider me putting on  
10 (.) back in the team,

## 3.3 Test-taker perceptions and feedback

The use of video and establishing a clear audience for the IC and PC tasks were important features of task design, which the results of the survey and interviews indicated were successful. Since no marked between-group differences were identified across the three proficiency groups and two L1 groups, only overall results are presented here. Only selected results are reported in this section: the full results of the survey are presented in Appendix C.

### 3.3.1 IC tasks

Of the test-takers, 93.8% indicated that they knew who they were communicating with (Q10). More than half of them attributed the clear awareness of the audience to the use of video for IC tasks (54.2% [Q12]), and 93.8% of test-takers regarded the use of video either 'helpful' or 'very helpful' in understanding Jan and Dan's points (Q11). In the subsequent interviews, test-takers made positive comments about the use of video as enhancing the authenticity of communication, for example:

*The video makes me feel it's very approachable... This is better than just looking at the description [on the screen], because that's an actual person talking to you, and he is talking quite naturally, which of course makes it more authentic. I really like it, [as the test becomes] less boring. (C04, B2)*

*It was definitely better with videos so if you really see it on the visuals yes it was just a bit more human and realistic... these were really good, they were very easy to understand, and the videos moved well and it was not like they got stuck. (A35, B2)*

Moreover, the authenticity of the required interactional situations was supported by the survey and interview responses. The IC tasks involved attempting to persuade the listener to change their view and were designed to elicit IC features including acknowledging the partner's point of view, putting forward a differing point of view and providing justification. Of the test-takers, 89.6% confirmed that trying to persuade is a common occurrence (Q13), happening either 'sometimes' (47.9%) or 'often' (41.7%) in real life. Test-takers commented in the interview, *I think it happens at any time. I think it's quite common in our daily life (C08, B1)* and *I think that happens on a daily basis at the university or when you go out or something (A35, B2).*



In addition, the IC tasks included a request from the listener for clarification and/or elaboration in response to *I'm sorry, I didn't quite get the last point. Can you say that again, perhaps using different words or giving an example?* from Jan and Dan. The artificial wording of this request-prompt, which was intended to provide some scaffolding to construct a response, attracted some criticisms and suggestions for potential alternatives (e.g., *I've never been asked [to reiterate] this way...just simply, shorter and a bit rougher [prompt is better] maybe?* (A27, C); *'Sorry?' or 'What did you say again?' or something* (A39, B2)). However, the very act of responding to a follow-up question was regarded as authentic and usual occurrence:

*That often happens. People definitely won't agree with whatever you said to them...  
It's normal. I think it's a little perfunctory if they don't ask questions.* (C04, B2)

### 3.3.2 PC tasks

The survey and interview responses evidenced the test-takers' awareness of the social and power relationship between interlocutors and imposition, which were part of the intended construct of the PC tasks. Similar to the IC tasks, test-takers had clear ideas about the intended audience – 95.8% indicated they clearly knew who they were communicating with (Q19). However, regarding the power relationship, there were some varied perceptions. (i.e., only 50.0% of the test-takers said Sue was 'equally powerful' and 72.9% said Mr Swift (coach) was 'more powerful', as intended). While interview respondents reported considering 'the age' (A26, C) and 'the other person's status' (C24, B1) in constructing and adjusting their responses, the more varied perceptions of power relationship for PC(a) task may have come from other factors, such as the degree of imposition:

*I found it difficult to think...how I do it morally because she [Sue] has a killer argument that she has to work at 3 o'clock, and I was like, 'how?' So, I personally would not ask – I would postpone my doctor's appointment.* (A27, C)

Indeed, 72% reported that their request to Sue was 'difficult' or 'very difficult' for her to accept (Q21) and 81.2% reported that asking Sue for a favour was 'uncomfortable' or 'very uncomfortable' for them (Q22), while 43.8% and 52.1% found it '(very) difficult and (very) uncomfortable' to make a request to Mr Swift in PC(b) respectively (Q25, Q26). Nevertheless, the authenticity of the PC tasks was supported by the majority of test-takers who stated the interactional situations in question were common occurrences (i.e., asking a classmate for a favour: 72.9% (Q23); making a request to a teacher/coach: 62.5% (Q27)).



## 4. CONCLUSIONS

---

This research showed the potential to assess a wider speaking construct through the careful design of tasks that target specific features of IC and PC in ways explicitly designed to reflect the affordances of a semi-direct test. We employed software that did not require advanced technologies to which resource-limited examination providers may not have access, thus opening up the possibility for similar tasks to be created and used in a range of contexts. Findings point to the possibility for semi-direct speaking tasks to elicit a selected range of IC features, including *acknowledging an interlocutor's view*, *countering a view*, *persuading an interlocutor to change his/her mind*, and *clarifying/exemplifying a point* in order to resolve a breakdown in communication. Importantly, test-takers felt that they knew who they were talking to and this was evident in the way that speech was directed to the listener/interlocutor, including engaging with specific points in a way that was clearly intended to communicate with the listener. The test-taker responses on the PC tasks also demonstrated that these tasks were able to elicit a range of PC features, such as *projecting the upcoming request*, *including a problem statement*, and *face-related statements*. The use of the latter often signals an increasing awareness by the test-taker of the listener's potential reaction and could contribute to a positive outcome in the interaction.

We recognise the project had some limitations in terms of the number of participants, who came from only two countries. However, this was appropriate for a small-scale trial. We anticipate that the findings will have resonance for the broader language testing community. As noted earlier, the trade-off that is reflected in a computer-mediated speaking test is the foregrounding of practicality for the narrower construct that can be measured within limited types of task, compared to that of face-to-face speaking tests. Given the current status of ASR and SDS still finding it challenging to generate naturalistic, complex discourse between machines and humans (Litman et al., 2018), it is hoped that the exploration of new tasks by targeting specific features of IC and PC has shown a promising way to broaden the construct that can be elicited by a computer-delivered speaking test.

As lines of future research based on the current dataset, the microanalysis of responses to the prototype IC and PC tasks will enable the drafting of rating criteria and descriptors at each level of performance, as well as help us to evaluate how well each of the four proposed tasks function to discriminate between levels. Furthermore, although the current study categorised performances as B1, B2 and C on the basis of a monologic task that did not target IC or PC constructs, future research could rate test-taker responses on the IC and PC tasks as more/less effective in terms of the features the tasks were designed to elicit. This could provide a different perspective on levels and inform rating scales.

Effective task design features can also be further explored. We aimed to maximise the use of visual cues to enhance the authenticity of communication and to enrich contextual information, and the decision to use video-recorded prompts indeed enabled test-takers to feel that they knew with whom they were communicating. However, the amount of text on screen and how best to use video input still require further investigation. For example, although interactive listening was an IC feature that the tasks were designed to elicit, the provision of key points from the conversational partner on screen might have compromised the extent to which the interactive listening construct could be measured. Future studies might involve a comparison of responses to tasks where the key points are provided and parallel tasks where the test-taker has to rely on listening to understand these points while care is taken not to add unnecessary memorisation demands.

## REFERENCES

- Al-Gahtani, S., & Roever, C. (2012). Proficiency and sequential organisation of L2 requests. *Applied Linguistics*, 33(1), 42–65. <https://doi.org/10.1093/applin/amr031>
- Council of Europe. (2001). *A Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.
- Ficzere, E. (2019). *An investigation into assessing ESL learners pragmatic competence at B2–C2 levels*. Unpublished PhD dissertation. University of Bedfordshire, UK.
- Galaczi, E. D., & Taylor, L. (2018). Interactional competence: Conceptualisations, operationalisations, and outstanding questions. *Language Assessment Quarterly*, 15(3), 219–236. <https://doi.org/10.1080/15434303.2018.1453816>
- Ginther, A. (2002). Context and content visuals and performance on listening comprehension stimuli. *Language Testing*, 19(2), 133–167. <https://doi.org/10.1191/0265532202lt225oa>
- Ikeda, N. (2017). *Measuring L2 Oral Pragmatic Abilities for Use in Social Contexts: Development and Validation of an Assessment Instrument for L2 Pragmatics Performance in University Settings*. Unpublished PhD thesis. University of Melbourne, Australia
- Lam, D. M. K. (2018). What counts as “responding”? Contingency on previous speaker contribution as a feature of interactional competence. *Language Testing*, 35(3) 377–401. <https://doi.org/10.1177/0265532218758126>
- Lee, S. H. (2009). Extended requesting: Interaction and collaboration in the production and specification of requests. *Journal of Pragmatics*, 41(6), 1248–71. <https://doi.org/10.1016/j.pragma.2008.09.013>
- Lim, G. S. (2018). Conceptualising and Operationalising Second Language Speaking Assessment: Updating the Construct for a New Century. *Language Assessment Quarterly*, 15(3), 215–218. <https://doi.org/10.1080/15434303.2018.1482493>
- Litman, D., Young, S., Gales, M., Knill, K., Ottewell, K., van Dalen, R., & Vandyke, D. (2016). Towards using conversations with spoken dialogue systems in the automated assessment of non-native speakers of English, *Proceedings of the SIGDIAL 2016 Conference*. Los Angeles, USA. <https://www.aclweb.org/anthology/W16-3635.pdf>
- Litman, D., Strik, H., & Lim, G. (2018). Speech Technologies and the Assessment of Second Language Speaking: Approaches, Challenges and Opportunities. *Language Assessment Quarterly*, 13(3), 294–309. <https://doi.org/10.1080/15434303.2018.1472265>
- May, L., Nakatsuhara, F., Lam, D., & Galaczi, E. (2020). Developing tools for learning oriented assessment of interactional competence: Bridging theory and practice. *Language Testing*, 37(2), 165–188. <https://doi.org/10.1177/0265532219879044>
- O’Sullivan, B., Dunlea, J., Spiby, R., Westbrook, C., & Dunn, K. (2020). *Aptis General Technical Manual Version 2.2* (TR/2020/001). London: British Council. <https://www.britishcouncil.org/exam/aptis/research/publications/technical/general-technical-manual-version-2-2>

- Plough I, Banerjee J., & Iwashita N. (2018). Interactional competence: Genie out of the bottle. *Language Testing*, 35(3): 427–445. <https://doi.org/10.1177%2F0265532218772325>
- Roever, C., & Kasper, G. (2018). Speaking in turns and sequences: Interactional competence as a target construct in testing speaking. *Language Testing*, 35(3) 331–355. <https://doi.org/10.1177%2F0265532218758128>
- Ross, S. (2018). Listener response as a facet of interactional competence. *Language Testing*, 35(3), 357–375. <https://doi.org/10.1177/0265532218758125>
- Taguchi, N., & Roever, C. (2017). *Second language pragmatics*. New York: Oxford University Press.
- Taleghani-Nikazm, C., & Huth, T. (2010). L2 requests: preference structure in talk-in-interaction. *Multilingua*, 29(2), 185–202. <https://doi.org/10.1515/mult.2010.008>
- Taylor, L. (Ed.) (2011). *Examining Speaking: Research and Practice in Assessing Second Language Speaking*. Studies in Language Testing 30. Cambridge: Cambridge ESOL/Cambridge University Press.
- Timpe-Laughlin, V., Evanini, K., Green, A., Blood, I., Dombi, J., & Ramanarayanan, V. (2017). Designing interactive, automated dialogues for L2 pragmatics learning. *Proceedings of 21<sup>st</sup> Workshop on the Semantics and Pragmatics of Dialogue*.
- van Batenburg, E. S. L., Oostdam, R. J., van Gelderen, A. J. S., & de Jong, N. H. (2018). Measuring L2 speakers' interactional ability using interactive speech tasks. *Language Testing*, 35(1), 75–100. <https://doi.org/10.1177%2F0265532216679452>
- Weir, C. J. (2005). *Language testing and validation: An evidence-based approach*. Basingstoke: Palgrave Macmillan.
- Youn, S. J. (2015). Validity argument for assessing L2 pragmatics in interaction using mixed methods. *Language Testing*, 32(2), 199–225. <https://doi.org/10.1177%2F0265532214557113>
- Young, R. (2011). Interactional competence in language learning, teaching, and testing. In E. Hinkel (Ed.), *Handbook of research of second language teaching and learning* (pp. 436–448). Abingdon: Routledge.
- Young, R. (2019). Interactional Competence and L2 Pragmatics, In N. Taguchi (Ed.), *The Routledge handbook of second language acquisition and pragmatics* (pp. 93–110). New York: Routledge.

## Appendix A: IC and PC Task Specifications

TEST	Aptis Research Pilot Test	COMPONENT	Speaking	Task	IC (a)/(b)
Features of the Test					
Task focus	The task is designed to tap into four features of IC: <i>responding to a partner, negotiating towards a joint outcome, interactive listening and negotiating meaning</i> . More specifically, the task measures candidates' ability to: <ul style="list-style-type: none"><li>disagree and put forward a different point of view <b>effectively</b> and provide <b>justification</b></li><li>effectively <b>link</b> their own <b>contribution</b> to the <b>partner's</b></li><li>work towards a decision by trying to <b>persuade</b> the partner</li><li><b>acknowledge partner's views</b></li><li>demonstrate they <b>have been listening</b> carefully/attentively through responding appropriately to the partner's idea</li><li>be able to <b>clarify/rephrase</b> their points</li></ul>				
Task level (CEFR)	A1	A2	B1	B2	C1 C2
Task description	Candidates will first watch a video in which a conversational partner expresses views, and then will be asked to express a differing opinion while appropriately responding to and persuading the partner, also using interactive listening skills. Key words from the partner's points as well as the points that the candidates are required to make will appear on screen. After the completion of the response time, an automated video clip will be played to ask a clarification question, to which the candidate will be required to respond.				
Instructions to candidates	In this part, you will watch the video of [relationship (e.g. your classmate, friend), NAME], who will argue for [yyy]. [NAME] will make two points in her talk. After you watch [NAME]'s talk, you will be given the two points that you're going to make. You will have 60 seconds to prepare your argument and 75 seconds to speak. Please talk to [NAME] in a way that shows that you understand [NAME]'s points, and use the two points to try to persuade [NAME] to change [his/her] opinion. After you have finished speaking, you will hear a question from [NAME]. You will be given 30 seconds to answer the question.				
Presentation of rubric	Aural		Written		Visual/ non-verbal [video]
Response format	Q&A		Short turn	Long turn	Interaction [Simulated]
Overall timing	4.5 minutes (tbc)				
Planning time	1 minute				
Delivery	Face-to-face		Telephone	Computer	Other
Nature of input	Real time (face-to-face)		Real time (remonte)	Pre-recorded input	No aural input
Nature of interaction	Unscripted		Guided	Semi-scripted	Scripted N/A
	Interlocutor-Candidate (I-C)			Candidate-Candidate (C-C) [Simulated]	
	Candidate only (C)			Interlocutor-Candidate-Candidate (I-C-C)	
Features of the Input / Prompt					
Description	A conversational partner expresses his/her views on a specific topic. The views include two specific points. During the preparation time, (i) key words from the partner's points as well as (ii) the points that the candidates are required to make appear on screen. (i) The key words from the partner's talk should be limited to 1-2 words and should be phrased in the way that should help to recall the candidate's memories without disclosing the partner's views to those key words. (ii) The two points that the candidates are going to make will also be aurally presented. Given the difficulty to switch between aural and written input modes, written information on the screen should be of a supportive nature throughout the task. After the completion of the response time (i.e. 75 seconds), an automated video clip is played to ask a clarification question, to which the candidate is required to respond.				
Video prompt	A conversational partner's talk: Between 140-160 words; No key information in the first 30 words; Start with some contextual information and then share two main points; The prompt to finish with 'What do you think?' There will be a standardised clarification request: "Sorry, I didn't get the last point. Can you say that again, perhaps using different words or giving an example?"				
Length of written prompt	Two bullet points for a conversational partner's views. Two bullet points for the candidate's views. Each of the candidate's bullet point is no longer than 12 words.				
Lexical level	K1	K2	K3	K4	K5 K6 K7 K8 K9 K10 TBC
Grammatical range	Most lexical items up to B1 according to the English Vocabulary Profile list				
Content knowledge	Most grammatical structures up to B1 according to the English Grammar Profile list				
Cultural specificity	General Neutral Specific Specific				

Nature of information	Only concrete		Mostly concrete		Fairly abstract		Mainly abstract
Relevant domain	Public		Occupational		Educational		Personal
Topic	From topic list for B2						
Topic familiarity	Familiar						Unfamiliar
Discourse mode	Descriptive	Biographical	Expository	Argumentative	Instructive	[quasi-] interactive	
Presentation	Verbal		Non-verbal (i.e. graphs)			Both	
Interlocutor	Speech rate: 160-170 wpm (tbc)						
	Variety of accent: Standard British English						
	Number of speakers: 2						
	Acquaintanceship: Friends (distance to be defined later)						
	Gender: Not specified						
Features of the Expected Response(s) – Spoken Output							
Description	Candidate will express a differing opinion with two points that are specified on screen, while appropriately acknowledging a conversational partner's views and trying to persuade the partner. Upon completion of the response time (75 seconds), an automated video clip will be played to ask a clarification question, to which the candidate will respond.						
Discourse mode	Descriptive	Biographical	Expository	Argumentative	Instructive	[quasi] interactive	
Presentation	Verbal		Non-verbal (i.e. graphs)			Both	
Length of response	75 seconds + 30 seconds						
Lexical level	Up to B2 lexis sufficient to respond adequately to task requirements (Suggestions to be made after the research)						
Structural range	Demonstration of grammatical control/range at the B2 level necessary for a rating of 3 (out of 5) for the task. (Suggestions to be made after the research)						
Functions targeted	Informational functions		Interactional functions		Managing interaction		
	Providing personal information		Agreeing		Initiating		
	Expressing opinions/preferences		Disagreeing		Changing topics		
	Elaborating		Modifying/commenting		Reciprocating		
	Justifying opinions		Asking for opinions		Deciding		
	Comparing		Persuading				
	Speculating		Asking for information				
	Staging		Conversational repair				
	Describing		Negotiation of meaning				
	Summarising						
	Suggesting						
Rating scale for task	A task-specific holistic rating scale is used for the task. The rating scale is a 7-point scale from 0–6. A B2-level performance is required to achieve score bands 3–4. A score of 5 or 6 is awarded for performances beyond B2 level, with a 5 describing performance equivalent to a C1 level, and 6 for performances at a C2 level. (Suggestions for descriptors and rating methods to be made after the research)						



TEST	Aptis Research Pilot Test		COMPONENT	Speaking	Task	PC (a)/(b)
Features of the Test						
Skills focus	<p>The task is designed to measure candidates' pragmatic competence in making requests with a relatively high level of imposition. The task targets the candidate's ability to:</p> <ul style="list-style-type: none"><li>• structure a coherent sequence, including logical pre-expansion features such as projecting the upcoming request and providing a reason for the request, which achieves the communicative goal of 'requesting'</li><li>• connect the different pre-expansion features appropriately (e.g. using conjunctions) in order to aid the listener's comprehension of the entire communicative act</li><li>• deliver the intended meanings while being sensitive to the social and power relationship between interlocutors, imposition and showing awareness of these contextual factors by using language for mitigation</li><li>• use a wide range of pragmalinguistic devices (e.g. downtoners, hedges, conversational routines) appropriately and in accordance with the level of imposition</li></ul>					
Task level (CEFR)	A1	A2	B1	B2	C1	C2
Task description	<p>The candidate will be given a situation to make a request (n.b. An element of the speech act of 'apology' may also be included). The situation prompts the candidate to recognise the pragmatic demands underlying the task and to initiate, develop and close his/her sequence to achieve the communicative goal of 'request'. The task provides a reason for the request to minimise cognitive demands related to <i>conceptualisation</i>, and the reason is legitimate or relatively legitimate. The task might request the hearer to perform a wish (e.g. cleaning the kitchen) or request the hearer to grant a wish (e.g. granting an essay extension).</p> <p>In PC(a), the power relationship between the speaker and hearer is equal, while that of PC (b) is unequal. To enrich contextual information, the task instructions include a very short video clip of the hearer opening a conversation.</p>					
Instructions to candidates	<p>In this part, you will be given a situation to make a request appropriately. After a situation is explained, you will have 60 seconds to prepare your talk. You will then have 60 seconds to speak.</p> <p>[Before preparation time] You have 60 seconds to prepare your answer. After 60 seconds, you will see the same video clip of [NAME]. After [NAME]'s greeting and question, you have 60 seconds to speak.</p>					
Presentation of rubric	Aural		Written		Visual/ non-verbal [video]	
Response format	Q&A		Short turn		Long turn	
Overall timing	3 minutes (tbc)					
Planning time	30 seconds					
Delivery	Face-to-face		Telephone		Computer	
Nature of input	Real time (face-to-face)		Real time (remonte)		Pre-recorded input	
	Unscripted		Guided		Semi-scripted	
					Scripted	
Nature of interaction	Interlocutor-Candidate (I-C) [Simulated role-play]				Candidate-Candidate (C-C) [Simulated role-play]	
	Candidate only (C)				Interlocutor-Candidate-Candidate (I-C-C)	
Features of the Input / Prompt						
Description	<p>Following task instructions, audio and written prompts will be given to the candidate, explaining a situation where he/she is asked to make a request. It should be noted that aural instructions are considered as the main source of information, and written instructions play a supportive role. This is to avoid increasing candidates' cognitive demands to switch between aural/spoken and written modes of communication. The prompt clearly provides background information and a legitimate (or relatively legitimate) reason for making the request.</p> <p>In PC(a), the power relationship between the speaker and hearer is equal, while that of PC (b) is unequal. The task instructions include a very short video clip of the hearer opening a conversation.</p>					
Video prompt	<p>The video prompt opens the conversation in PC (a) with 'Hi, is everything OK? You look a bit worried.' and opens the conversation in PC(b) with 'Hello. How can I help you?'</p>					

Length of written prompt	The length of a prompt (excluding task instructions regarding preparation and response times) ranges between 120-140 words.										
Lexical level	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	TBC
Grammatical range	Most lexical items up to B1 according to the English Vocabulary Profile list										
Content knowledge	Most grammatical structures up to B1 according to the English Grammar Profile list										
Cultural specificity	General							Specific			
Nature of information	Neutral							Specific			
Nature of information	Only concrete			Mostly concrete			Fairly abstract			Mainly abstract	
Relevant domain	Public			Occupational			Educational			Personal	
Topic	From topic list for B2										
Topic familiarity	Familiar							Unfamiliar			
Discourse mode	Descriptive		Biographical		Expository		Argumentative		Instructive		[quasi-] Interactive
Presentation	Verbal				Non-verbal (i.e. graphs)				Both		
Interlocutor	Speech rate: 160-170 wpm (tbc)										
	Variety of accent: Standard British English										
	Number of speakers: 2										
	Acquaintanceship: PC(a) classmate/colleague (equal power relationship); PC(b) authority figure (unequal power relationship)										
	Gender: Not specified (either same or opposite gender)										
Features of the Expected Response(s) – Spoken Output											
Description											
Discourse mode	Descriptive		Biographical		Expository		Argumentative		Instructive		[quasi] interactive
Presentation	Verbal				Non-verbal (i.e. graphs)				Both		
Length of response	1 minute										
Lexical level	Up to B2 lexis sufficient to respond adequately to task requirements (Suggestions to be made after the research)										
Structural range	Demonstration of grammatical control/range at the B2 level necessary for a rating of 3 (out of 5) for the task. (Suggestions to be made after the research)										
Functions targeted	Informational functions				Interactional functions				Managing interaction		
	Providing personal information				Agreeing				Initiating		
	Expressing opinions/preferences				Disagreeing				Changing topics		
	Elaborating				Modifying/commenting				Reciprocating		
	Justifying opinions				Asking for opinions				Deciding		
	Comparing				Persuading						
	Speculating				Asking for information						
	Staging				Conversational repair						
	Describing				Negotiation of meaning						
	Summarising				Requesting						
	Suggesting										
Rating scale for task	A task-specific holistic rating scale is used for the task. The rating scale is a 7-point scale from 0–6. A B2-level performance is required to achieve score bands 3–4. A score of 5 or 6 is awarded for performances beyond B2 level, with a 5 describing performance equivalent to a C1 level, and 6 for performances at a C2 level. (Suggestions for descriptors and rating methods to be made after the research)										

## Appendix B: Task instructions and prompts

### **IC (a) Task instructions and prompts**

#### **Screen 1: The same instructions are displayed as the aural instructions**

In this part, you will watch the video of your classmate Jan, who will argue for **the use of group presentations**.

**Jan will make 2 points** in her talk.

After you watch Jan's talk, you will be given **the 2 points that you're going to make**. You will have:

- **60 seconds to prepare** your argument, and
- **75 seconds to speak**.

Please talk to Jan in a way that **shows that you understand Jan's points**, and **use the two points to try to persuade** Jan to change her opinion.

After you have finished speaking, you will hear **a question from Jan**. You will be given **30 seconds to answer** the question.

### **Aural instructions for Screen 1 (Transcripts)**

In this part, you will watch the video of your classmate Jan, who will argue for **the use of group presentations**. **Jan will make 2 points** in her talk.

After you watch Jan's talk, you will be given **the 2 points that you're going to make**. You will have **60 seconds to prepare** your argument and **75 seconds to speak**.

Please talk to Jan in a way that **shows that you understand Jan's points**, and **use the two points to try to persuade** Jan to change her opinion.

After you have finished speaking, you will hear **a question from Jan**. You will be given **30 seconds to answer** the question.

## Screen 2

In a classroom, Jan starts talking to you.

[Jan's video – In a classroom of a college/university, Jan starts talking to you. Jan's upper body shown with a table. Some hand gestures and facial expressions.

### Aural prompt for Screen 2 (Transcripts)

In a classroom, Jan starts talking to you.

Have you heard that we can now vote on whether to do a group or individual presentation next term? I'm definitely going to vote for doing a group presentation, and I hope you will, too. I think group presentations are really great and they are so much better than individual presentations. In my experience, working together with friends develops our teamwork skills. You know, I'm so pleased that I learnt how to become a good team member during the last group project, and it was wonderful that all of us worked hard for the team. And... I also believe working in groups is more similar to our real-life activities in the future. When we start working, we'll need to work together with our colleagues. So... to me, it makes more sense to do group presentation tasks now to practise what we'll be doing in the future. What do you think?



### Screen 3

Use the two points below to persuade Jan to change her opinion. Remember to first show your understanding of her points.

Jan's points:

- Teamwork skills
- Real-life

**Your points:**

- **Group presentations only work if all team members work hard**
- **Probably unfair, if all team members receive the same score**

You now have 60 seconds to think about your answer to Jan.

**Timer (60 seconds)**

### Aural instructions for Screen 3 (Transcripts)

In your talk, use the two points on the screen to persuade Jan to change her opinion. The two points that you're going to make are 'Group presentations only work if all team members work hard' and 'It is probably unfair, if all team members receive the same score'. Key words from Jan's points are also provided.

Remember to first show your understanding of her points. You now have 60 seconds to think about your answer to Jan.

## Screen 4

You now have 75 seconds to talk to Jan, making two points to persuade her to change her opinion. Remember to show your understanding of her points.

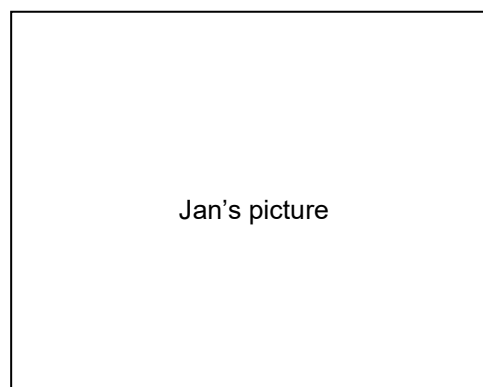
Jan's points:

- Teamwork skills
- Real-life

**Your points:**

- **Only works if all team members work hard**
- **Probably unfair, if all team members receive the same score**

**Timer 75 seconds)**



## Aural instructions for Screen 4 (Transcripts)

You now have 90 seconds to talk to Jan, making two points to persuade her to change her opinion. Remember to show your understanding of her points. Please start speaking after the Beep.

[Beep]

## Screen 5

[Jan's video]

Now Jan has a question for you.

You have 30 seconds to answer after the Beep.

**Timer (30 seconds)**

## Aural prompt for Screen 5 (Transcripts)

Now Jan has a question for you. You have 30 seconds to answer after the beep.

Sorry, I didn't quite get the last point. Can you say that again, perhaps using different words or giving an example?

[Beep]

## **IC (b) Task instructions and prompts**

### **Screen 1: The same instructions are displayed as the aural instructions**

In this part, you will watch the video of your classmate Dan, who will argue for **travelling alone**.

**Dan will make 2 points** in his talk.

After you watch Dan's talk, you will be given **the 2 points that you're going to make**. You will have:

- **60 seconds to prepare** your argument, and
- **75 seconds to speak**.

Please talk to Dan in a way that **shows that you understand Dan's points**, and **use the two points to try to persuade** Dan to change his opinion.

After you have finished speaking, you will hear **a question from Dan**. You will be given **30 seconds to answer** the question.

### **Aural instructions for Screen 1 (Transcripts)**

In this part, you will watch the video of your classmate Dan, who will argue for travelling alone. Dan will make two points in his talk.

After you watch Dan's talk, you will be given the two points that you're going to make. You will have 60 seconds to prepare your argument and 75 seconds to speak.

Please talk to Dan in a way that **shows that you understand Dan's points**, and **use the two points to try to persuade** Dan to change his opinion.

After you have finished speaking, you will hear a question from Dan. You will be given 30 seconds to answer the question.

## Screen 2

In a cafe, Dan starts talking to you.

[Dan's video] In a cafe, Dan starts talking to you. Dan's upper body shown with a table. Some hand gestures and facial expressions.

### Aural prompt for Screen 2 (Transcripts)

In a cafe, Dan starts talking to you.

Hi, do you have a few minutes? I'd really like to hear your opinion. Do you remember that I was gonna go to Scotland on my own this summer? But when I was talking about my holiday plans with my friends yesterday, they said that they'd like to join me for the trip. They're good friends so don't get me wrong, but... I'm actually thinking of telling them that I'd like to travel alone this time. I think travelling alone is better than travelling with friends. I can change my plans whenever I want to. I don't need to discuss what time to wake up, what or where to eat, or whether to see a museum or go hiking. And... the other issue is if you travel alone, you're more likely to meet local people and experience local culture. What do you think?

## Screen 3

Use the two points below to persuade Dan to change his opinion. Remember to first show your understanding of his points.

Dan's points:

- Change plans
- Local people

Your points:

- Can share experiences with friends
- Can be cheaper to share costs with friends

You now have 60 seconds to prepare your answer to Dan.

**Timer (60 seconds)**

### Aural instructions for Screen 3 (Transcripts)

In your talk, use the two points below to persuade Dan to change his opinion. The two points that you're going to make are 'You can share experiences with friends' and 'it can be cheaper to share costs with friends'. Remember to first show your understanding of his points. Key words from Dan's talk are also provided. You now have 60 seconds to think about your answer to Dan.

## Screen 4

You now have 75 seconds to talk to Dan, making two points to persuade him to change his opinion. Remember to show your understanding of his points.

Dan's points:

- Change plans
- Local people

**Your points:**

- Can share experiences with friends
- Can be cheaper to share costs with friends

**Timer (90 seconds)**

Dan's picture

## Aural instructions for Screen 4 (Transcripts)

You now have 75 seconds to talk to Dan, making two points to persuade him to change his opinion. Remember to show your understanding of his points. Please start speaking after the Beep.

[Beep]

## Screen 5

[Dan's video]

Now Dan has a question for you.

You have 30 seconds to speak after the Beep.

**Timer (30 seconds)**

## Aural prompt for Screen 5 (Transcripts)

Now Dan has a question for you. You will have 30 seconds to answer after the beep.

Sorry, I didn't quite get the last point. Can you say that again, perhaps using different words or an example?

[Beep]

## **PC (a) Task instructions and prompts S=H (equal power)**

### **Screen 1**

In this part, you will be given a situation to make a request appropriately. After a situation is explained, you will have:

- 60 seconds to prepare your talk.
- 60 seconds to speak.

### **Aural instructions for Screen 1 (Transcripts)**

In this part, you will be given a situation to make a request appropriately. After a situation is explained, you will have 60 seconds to prepare your talk. You will then have 60 seconds to speak.

### **Screen 2: A video clip to appear in the second paragraph (Two instruction paragraphs to appear together with aural prompt)**

[A still picture of Sue] You are going to do a presentation for your class this week. You agreed to present at 3pm tomorrow, but you now need to go and see a doctor at that time. You see that your classmate Sue is going to present at 1pm. So, you decide to ask Sue if she could exchange times with you. It might not be easy for Sue, as she usually starts at a part-time job at 3pm.

As you're walking outside the classroom, you see Sue. She stops and talks to you.

[To play a video clip of Sue]<sup>2</sup> "Hi, is everything OK? You look a bit worried."

Please ask Sue to exchange her presentation time with you. Remember to explain your situation and also to consider Sue's situation.

**You have 60 seconds to prepare your answer and 60 seconds to speak.**

**Please start to prepare your answer after the beep.**

**Timer (60 seconds)**

A video clip of Sue

### **Aural prompt for Screen 2 (Transcripts)**

You are going to do a presentation for your class this week. You agreed to present at 3pm tomorrow, but you now need to go and see a doctor at that time. You see that your classmate Sue is going to present at 1pm. So, you decide to ask Sue if she could exchange times with you. It might not be easy for Sue, as she usually starts at a part-time job at 3pm.

---

<sup>2</sup> Not shown on screen

As you're walking outside the classroom, you see Sue. She stops and talks to you.

Please ask Sue to exchange her presentation time with you. Remember to explain your situation and also to consider Sue's situation.

You have 60 seconds to prepare your answer and 60 seconds to speak.

Please start to prepare your answer after the beep.

[Beep]

### Screen 3: A video and a response timer appear

You are going to do a presentation for your class this week. You agreed to present at 3pm tomorrow, but you now need to go and see a doctor at that time. You see that your classmate Sue is going to present at 1pm. So, you decide to ask Sue if she could exchange times with you. It might not be easy for Sue, as she usually starts at a part-time job at 3pm.

As you're walking outside the classroom, you see Sue. She stops and talks to you.

[To play a video clip of Sue]<sup>3</sup> "Hi, is everything OK? You look a bit worried."

Please ask Sue to exchange her presentation time with you. Remember to explain your situation and also to consider Sue's situation.

**Please start speaking after Sue talks to you.**

**Timer (60 seconds)**

A video clip of Sue

### Aural prompt for Screen 3 (Transcripts)

Now, please start speaking after Sue talks to you.

[A video clip of Sue appears] Sue: "Hi, is everything OK? You look a bit worried."

[Beep]

---

<sup>3</sup> Not shown on screen



## **PC (b) Task instructions and prompts Request S<H (unequal power)**

### **Screen 1**

In this part, you will be given a situation to make a request appropriately. After a situation is explained, you will have:

- 60 seconds to prepare your talk
- 60 seconds to speak.

### **Aural instructions for Screen 1 (Transcripts)**

In this part, you will be given a situation to make a request appropriately. After a situation is explained, you will have 60 seconds to prepare your talk. You will then have 60 seconds to speak.

### **Screen 2: A video clip to appear in the second paragraph (Two instruction paragraphs to appear together with aural prompt)**

[A still picture of Mr Swift] You were selected for the school basketball team, and played your first game last week. You've practiced hard to prepare, but on the day of the game you had a bad headache. You didn't tell anyone because you didn't want to let the team down. As a result, you didn't play well, and the coach has now dropped you from the team. You decide to ask your coach, Mr Swift, to give you another chance. But you know that this might be unfair to other teammates.

You knock on Mr Swift's office door and you're asked to enter. He says, [To play a video clip of Mr Swift]<sup>4</sup> "Hello. How can I help you?"

Please ask Mr Swift to consider putting you back in the team. Remember to explain your situation and consider Mr Swift's situation.

**You have 60 seconds to prepare your answer and 60 seconds to speak.**

**Please start to prepare your answer after the beep. [Beep]**

**Timer (60 seconds)**

A video clip of Mr Swift

### **Aural prompt for Screen 2 (Transcripts)**

You were selected for the school basketball team, and played your first game last week. You've practiced hard to prepare, but on the day of the game you had a bad headache. You didn't tell anyone because you didn't want to let the team down. As a result, you didn't play well, and the coach has now dropped you from the team. You decide to ask your coach, Mr Swift, to give you another chance. But you know that this might be unfair to other teammates.

<sup>4</sup> Not shown on screen

You knock on Mr Swift's office door and you're asked to enter. He says, [To play a video clip of Mr Swift] "Hello. How can I help you?"

Please ask Mr Swift to consider putting you back in the team. Remember to explain your situation and consider Mr Swift's situation.

You have 60 seconds to prepare your answer and 60 seconds to speak. Please start to prepare your answer after the beep.

[Beep]

### Screen 3: The last paragraph and a preparaton timer appear

[A still picture of Mr Swift] You were selected for the school basketball team, and played your first game last week. You've practiced hard to prepare, but on the day of the game you had a bad headache. You didn't tell anyone because you didn't want to let the team down. As a result, you didn't play well, and the coach has now dropped you from the team. You decide to ask your coach, Mr Swift, to give you another chance. But you know that this might be unfair to other teammates.

You knock on Mr Swift's office door and you're asked to enter. He says, [To play a video clip of Mr Swift]<sup>5</sup> "Hello. How can I help you?"

Please ask Mr Swift to consider putting you back in the team. Remember to explain your situation and consider Mr Swift's situation.

**Now, please start speaking after Mr Swift talks to you.**

**Timer (60 seconds)**

A video clip of Mr Swift

### Aural prompt for Screen 4 (Transcripts)

Now, please start speaking after Mr Swift talks to you.

[A video clip of Mr Swift appears] "Hello. How can I help you?"

[Beep]

---

<sup>5</sup> Not shown on screen

## Appendix C: Post-test survey results

Thank you for participating in the speaking test. We'd like to know a bit **about you** and **how you felt about the test tasks**!

### About you

1. My name is [ \_\_\_\_\_ ]

2. My candidate ID number is [ \_\_\_\_\_ ]

3. My first language (mother tongue) is: (please tick)

	[TOTAL (N=48)]	Chinese (n=24)	Austrian (n=24)]
a. Mandarin	[50.0%	100.0%	0.0%]
b. Cantonese	[0.0%]		
c. German	[45.8%	0.0%	91.7%]
d. Other (please specify)	Croatian and Turkish [2.1% each against TOTAL; 4.2% against Austrian]		

4. I am ..... years old.

	[TOTAL (N=48)]	Chinese (n=24)	Austrian (n=24)]
a. under 16 .....	[0.0%]		
b. 17 .....	[0.0%]		
c. 18 .....	[18.8%	20.8%	16.7%]
d. 19 .....	[31.3%	37.5%	25.0%]
e. 20 .....	[20.8%	37.5%	4.2%]
f. 21 .....	[6.3%	0.0%	12.5%]
g. 22 .....	[4.2%	0.0%	8.3%]
h. 23 .....	[4.2%	4.2%	4.2%]
i. 24 .....	[8.3%	0.0%	16.7%]
j. 25 .....	[0.0%]		
k. 26 .....	[0.0%]		
l. 27 .....	[2.1%	0.0%	4.2%]
m. 28 .....	[0.0%]		
n. 29 .....	[2.1%	0.0%	4.2%]
o. 30 .....	[2.1%	0.0%	4.2%]
p. over 30 .....	[0.0%]		
q. Prefer not to say .....	[0.0%]		

5. I am:

a. male	[56.3%	50.0%	62.5%]
b. female	[39.6%	45.8%	33.3%]
c. Prefer not to say	[4.2%	4.2%	4.2%]

6. I am:

a. a student at high school	[18.7%	0.0%	37.5%]
b. a student at college/university	[81.3%	100.0%	62.5%]
c. working full time	[0.0%		
d. Other (please specify)	[0.0%		

## About the speaking tasks

**Section 1: We'd like you to think about the two tasks where you tried to persuade Jan and Dan to change their view.**

<i>Photo</i>	<i>Photo</i>
<i>Jan</i>	<i>Dan</i>

[TOTAL (N=48)    Chinese (n=24)    Austrian (n=24)]

7. The time I had to prepare to answer Jan and Dan (60 seconds) was:

a. too short	[14.6%	16.7%	12.5%]
b. ok	[70.8%	79.2%	62.5%]
c. too long	[10.4%	0.0%	20.8%]
*missing data	[4.2%	4.2%	4.2%]

8. The time I had to speak to Jan and Dan (75 seconds) was:

a. too short	[22.9%	37.5%	8.3%]
b. ok	[58.3%	50.0%	66.7%]
c. too long	[16.7%	8.3%	25.0%]
*missing data	[2.1%	4.2%	0.0%]

9. The instructions for Jan and Dan's tasks were:

a. very clear	[52.1%	41.7%	62.5%]
b. clear	[43.8%	54.2%	33.3%]
c. not clear	[2.1%	0.0%	4.2%]
*missing data	[2.1%	4.2%	0.0%]

10. I knew who I had to communicate with in these tasks (i.e. Jan and Dan).

a. Yes	[93.8%	91.7%	95.8%]
b. Not sure	[4.2%	4.2%	4.2%]
c. No	[0.0%		
*missing data	[2.1%	4.2%	0.0%]

11. To understand Jan and Dan's points, their videos were:

a. very helpful	[52.1%	41.7%	62.5%]
b. helpful	[41.7%	50.0%	33.3%]
c. not helpful	[4.2%	4.2%	4.2%]
*missing data	[2.1%	4.2%	0.0%]

12. The video helped me to feel that I was communicating with Jan and Dan.

a. Yes	[54.2%	50.0%	58.3%]
b. Not sure	[29.2%	33.3%	25.0%]
c. No	[14.6%	12.5%	16.7%]
*missing data	[2.1%	4.2%	0.0%]

13. Trying to persuade someone to change their view would ..... happen in real life.

a. often	[41.7%	45.8%	37.5%]
b. sometimes	[47.9%	37.5%	58.3%]
c. rarely	[8.3%	12.5%	4.2%]
d. never	[2.1%	4.2%	0.0%]

14. Answering Jan and Dan's follow-up questions was:

a. easy	[27.1%	33.3%	20.8%]
b. ok	[52.1%	41.7%	62.5%]
c. difficult	[20.8%	25.0%	16.7%]

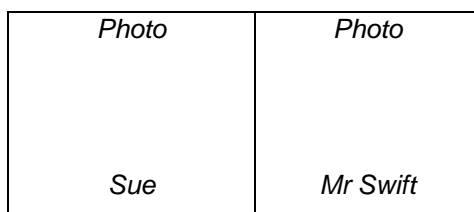
15. The time I had to answer Jan and Dan's follow-up questions (30 seconds) was:

a. too short	[27.1%	25.0%	29.2%]
b. ok	[66.7%	70.8%	62.5%]
c. too long	[6.3%	4.2%	8.3%]

16. Any other comments or feedback on Jan and Dan's tasks?

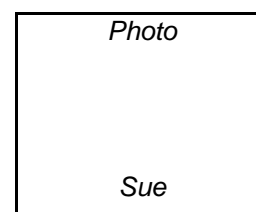
- I found the interaction very 'inhuman', e.g. 'Please can you say that in different words?'. I also don't express myself 'differently' or 'genählt' in German.
- I prefer American English accent.
- I thought I had 30 seconds to think about the answer after the questions from Jan and for that reason, I missed the chance to answers.
- It would be good to have a notepad for key words.
- No, everything was understandable.
- Quite good.
- Quite good. However, the time given to answer the follow-up question was shorter than the time given to answer the previous question. I could have answered more.
- The option to answer spontaneously was missing -> skip the 60 seconds of preparation
- The situations is quite realistic and close to our daily life. More time should be given to prepare for the topics on persuading people to change their perspectives.
- The time limit was too short for persuasion. I was only able to state my opinion but I don't have time to talk about the reasons in detail.
- Their (Jan & Dan's) statements were not clear enough, so it took me some time to organise the content.
- Very good.
- Very good. Very straightforward.
- Very novel.
- You expect – like with Sue and Mr Swift – that they repeat the input once more before you have to start speaking. With Jan and Dan, only the situation is introduced and the instructions followed by a 'bleep'.

**Section 2: We'd like you to think about the tasks where you made requests of your classmate Sue and your basketball coach Mr Swift.**



	[TOTAL(N=48)	Chinese (n=24)	Austrian (n=24)]
17. The time I had to prepare to talk to Sue and Mr Swift (60 seconds) was:			
a. too short	[20.8%	25.0%	16.7%]
b. ok	[72.9%	75.0%	70.8%]
c. too long	[6.3%	0.0%	12.5%]
18. The time I had to speak to Sue and Mr Swift (60 seconds) was:			
a. too short	[18.8%	16.7%	20.8%]
b. ok	[66.7%	75.0%	58.3%]
c. too long	[14.6%	8.3%	20.8%]
19. I knew who I had to communicate with in these tasks (i.e. Sue and Mr Swift).			
a. Yes	[95.8%	91.7%	100.0%]
b. Not sure	[4.2%	8.3%	0.0%]
c. No	[0.0%		

**Section 2.1: Now we'd like you to focus on the task where you asked your classmate Sue to swap presentation times with you.**



	[TOTAL (N=48)	Chinese (n=24)	Austrian (n=24)]
20. In this relationship, Sue was ..... me.			
a. more powerful than	[43.8%	41.7%	45.8%]
b. equally as powerful as	[50.0%	54.2%	45.8%]
c. less powerful than	[6.3%	4.2%	8.3%]



21. My request to Sue was ..... **for her** to do.

a. very easy	[4.2%	4.2%	4.2%]
b. easy	[22.9%	29.2%	16.7%]
c. difficult	[68.8%	66.7%	70.8%]
d. very difficult	[4.2%	0.0%	8.3%]

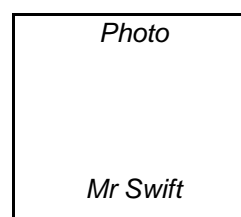
22. Asking Sue a favour was ..... **for me**.

a. very comfortable	[2.1%	4.2%	0.0%]
b. comfortable	[16.7%	20.8%	12.5%]
c. uncomfortable	[70.8%	62.5%	79.2%]
d. very uncomfortable	[10.4%	12.5%	8.3%]

23. Asking our classmates for a favour would ..... happen in real life.

a. often	[18.7%	12.5%	25.0%]
b. sometimes	[54.2%	58.3%	50.0%]
c. rarely	[27.1%	29.2%	25.0%]
d. never	[0.0%		

**Section 2.2: Finally, we'd like you to focus on the task where you asked your basketball coach Mr Swift to consider putting you back on the team.**



[TOTAL (N=48)      Chinese (n=24)      Austrian (n=24)]

24. In this relationship, Mr Swift was ..... me.

a. more powerful than	[72.9%	66.7%	79.2%]
b. equally as powerful as	[22.9%	33.3%	12.5%]
c. less powerful than me.	[4.2%	0.0%	8.3%]

25. My request to Mr Swift was ..... **for him** to do.

a. very easy	[12.5%	12.5%	12.5%]
b. easy	[43.8%	54.2%	33.3%]
c. difficult	[39.6%	25.0%	54.2%]
d. very difficult	[4.2%	8.3%	0.0%]

26. Making a request to Mr Swift was ..... **for me.**

a. very comfortable	[10.4%	16.7%	4.2%]
b. comfortable	[35.4%	33.3%	37.5%]
c. uncomfortable	[37.5%	37.5%	37.5%]
d. very uncomfortable	[14.6%	12.5%	16.7%]
*missing data	[2.1%	0.0%	4.2%]

27. Making a request to our teacher/coach would ..... happen in real life.

a. often	[16.7%	25.0%	8.3%]
b. sometimes	[45.8%	37.5%	54.2%]
c. rarely	[31.3%	33.3%	29.2%]
d. never	[6.3%	4.2%	8.3%]

28. Any other comments or feedback on Sue and Mr Swift's tasks?

- (The candidate responded in English) While talk to Sue it's not a common situation I meet in school. I need imagine Jan's situation by myself.
- Maybe the request given could be more reasonable, otherwise it would be hard to come up with the reasons to make the request.
- More than anything, the fact that other people in the room were talking disturbed me.
- Quite good.
- Speaking to Sue and getting her to do something that would be difficult for her was more difficult for me than speaking to Mr Swift.
- Sue's clothes?
- The background conditions of the requests can be relatively more flexible, which might be beneficial to my performance on this question.
- The given in the question makes me feel quite awkward, because for me this kind of conversation is quite unlikely to happen.
- The time for preparation is too short, and some of the requests are too difficult to be approved.
- Very good.
- Very good. The tasks that I needed to do were stated very clearly.

*Thank you very much!*

## Appendix D:

### Post-survey interview questions

---

#### **Instructions to interviewers/administrators.**

1. Please briefly look through the responses that the candidate has made to the survey items, so that you can **select from the following questions** those that are more relevant for each interviewee.
2. Interviews are to be conducted in **learners' L1**.
3. Please ensure that the interview is **audio-recorded** (file name to indicate Candidate ID).
4. Please translate and transcribe audio files.

*Thank you very much for agreeing to take part in this short interview. In this research we are interested in the way that you went about responding to the speaking tasks you have just completed and your thoughts on them.*

#### **IC tasks**

*I'd like you to think back to the tasks where you tried to persuade Jan and Dan to consider a different point of view.*

<i>Photo</i>	<i>Photo</i>
<i>Jan</i>	<i>Dan</i>

1. How did you feel about the **time** allocated for preparing and speaking on these tasks? [\[refer to answers to Q 7, 8, 15 if any indicate too short or too long\]](#).
2. What impact did the **use of video** have? [\[refer to answers to Q11 & 12: Depending on response, possible follow up Qs\]](#)
  - a. Did the video make the interaction seem more real? If yes, in what way?
  - b. Did it help you to understand the context? If yes, in what way?
  - c. What extra information did you get from the video that you would not have got from only written instructions? [Interviewer can hide the video window on screen/on paper by hand to demonstrate]
  - d. Do you prefer video or audio input? Why?

3. [For those who selected a or b in Q13] You said that trying to persuade a friend to change his/her opinion would (often/sometimes) happen **in real life**. In what situations might this happen?
4. Jan and Dan asked you a **follow-up question**.
  - a. Is this something that would happen in real life?
  - b. How did you feel about being asked a follow-up question?
  - c. Did you feel satisfied with your response to the follow-up question? Why? Why not?
5. The task required you to **show your understanding** of Jan's and Dan's points. How did you do this?
6. You commented [refer to comment box Q16]. Could you tell me more about this/explain why you felt this way?

### PC tasks

*I'd now like you to think back to the tasks where you asked Sue to change presentation times with you and Mr Swift to put you back on the basketball team.*

Photo	Photo
Sue	Mr Swift

7. How did you feel about the **time** allocated for preparing and speaking in these tasks? [refer to Q 17 & 18 if any indicate too short or too long].
8. [For those who ticked a or b in Q23] You said that asking a request to a classmate or friend would (often/sometimes) happen **in real life**. In what situations might this happen?
9. [For those who selected a or b in Q27] You said that asking a request to teacher, coach or someone in a leadership role would (often/sometimes) happen **in real life**. In what situations might this happen?
10. What did you consider when formulating your answers (e.g. the age of Sue/Mr Swift, language to use)?
  - a. Sue's task?
  - b. Mr Swift's task?
11. Do you think they will grant your wishes? Why/why not?
  - a. Sue's task?
  - b. Mr Swift's task?
12. You commented [refer to comment box Q28]. Could you please tell me more about this/explain why you felt this way?

## Appendix E: Transcription notations

<b>Unfilled pauses or gaps</b>	(.) for a short pause (less than approx. 1 second); (..) for a medium pause (approx. 1-3 seconds); (...) for a long pause (more than 3 seconds)
<b>Double parentheses ((laughter))</b>	Laughter is signalled by “laughter” in double parentheses
<b>Dash -</b>	A cut off
<b>Um, huh, eh, em</b>	Filled pauses within single parentheses as you hear them
<b>Empty parentheses ( )</b>	Words within single parentheses are doubtful or uncertain
<b>Square brackets [<i>nandaro</i>]</b>	Use square brackets for first language inserted into an utterance. Write an approximation of the sound inside the brackets. Italicise the word.
<b>Comma, period, question mark</b>	Use your judgement to add a comma, period and question mark as you hear the tone of the speaker.

# British Council Assessment Research Group

The Assessment Research Group was formed in 2013 to support the British Council's work in assessment and testing across the world. The team is responsible for ensuring that all new assessment products and new uses of existing products are supported by the most up-to-date research. They also continuously evaluate the quality of British Council assessment products.

## EXPLORING THE POTENTIAL FOR ASSESSING INTERACTIONAL AND PRAGMATIC COMPETENCE IN SEMI-DIRECT SPEAKING TESTS

VS/2021/001

**F. Nakatsuhara, L. May,  
C. Inoue, E. Willcox-Ficzere,  
C. Westbrook and R. Spiby**

### BRITISH COUNCIL VALIDATION SERIES

Published by British Council  
1 Redman Place, Stratford  
London E20 1JQ  
United Kingdom

© **British Council 2021**

The British Council is the  
United Kingdom's international  
organisation for cultural relations  
and educational opportunities.

[www.britishcouncil.org/aptis/research](http://www.britishcouncil.org/aptis/research)