Eliciting pragmatic and interactional competence in semi-direct speaking tests

Fumiyo Nakatsuhara & Lyn May

University of Bedfordshire

QUT Queensland University of Technology

Acknowledgements

- This project was funded and supported by the British Council.

- Other project members: Chihiro Inoue (CRELLA, University of Bedfordshire), Edit Willcox-Ficzere (Oxford Brookes), Carolyn Westbrook & Richard Spiby (British Council)

Nakatsuhara et al. (2021) *Exploring the potential for assessing interactional and pragmatic competence in semi-direct speaking tests,* Aptis Validation Report. British Council. https://www.britishcouncil.org/sites/default/files/nakatsuhara_et_al_2021_-_assessing_interactional_and_pragmatic_competence_in_semi-direct_speaking_tests_v2.pdf

BACKGROUND

How can we expand the construct that we assess in computer-delivered speaking tests?

- Issues in operationalising the IC construct (e.g. Galaczi & Taylor, 2018)

- Lack of authenticity in communicatively-oriented tasks (e.g. May, 2018)

"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: 232)

"Computer-based tests currently lack interactivity, which means that certain aspects of the IC construct cannot be operationalised"

(Plough et al. 2018: 439)

Possible solutions?

b) spoken dialogue systems

a) video-conferencing delivery

Trinity ISE Online
https://www.trinitycollege.it/inglese/esami-in-videoconferenza/

Matsuyama (2022)
https://www.teai-waseda.jp/assessment/
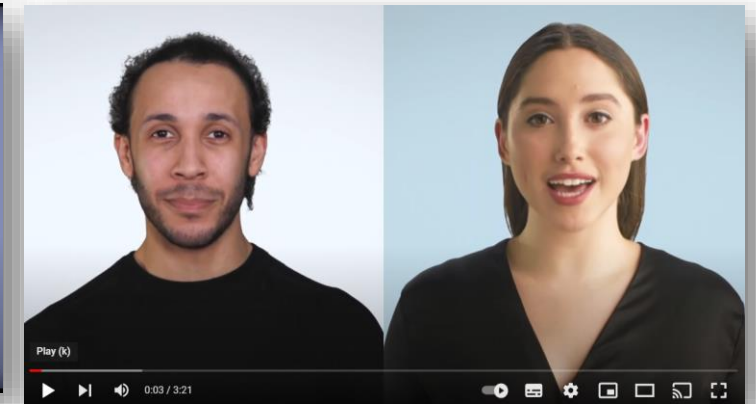
GPT-3: Two AIs talk about becoming human
https://www.youtube.com/watch?v=jz78fSnBG0s

d) virtual reality & augmented reality

e) carefully designed semi-direct tasks

c) virtual environments

(Ockey et al. 2017)

Mondly AR
https://www.youtube.com/watch?v=9P0t9JI22y8

Oxford Test of English
https://fdslive.oup.com/www.oup.com/elt/general_content/global/ote/demo-v3/#/exercise/speaking/2/1
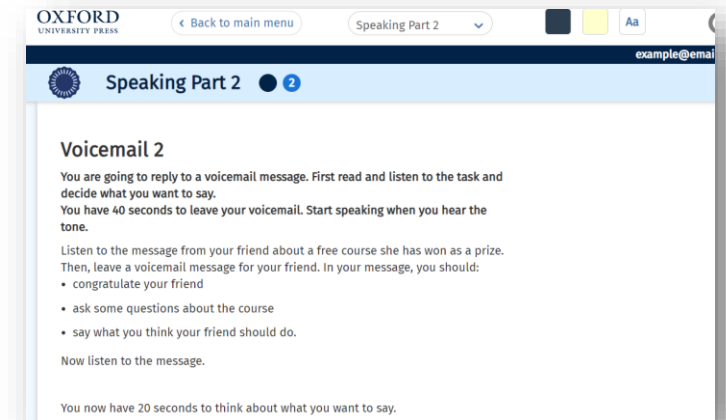
Research Questions

RQ1. Can a computer-based semi-direct speaking test elicit features of **pragmatic competence** at different levels?
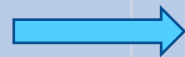
RQ2. Can a computer-based semi-direct speaking test elicit features of **interactional competence** at different levels?

METHODS

Overall research design: Data collection

Development of specifications for IC and PC tasks

Creation of 2 PC and 2 IC tasks → Piloting and revision of PC and IC tasks

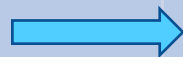**Trialling 1 benchmarking task + 2 PC tasks & 2 IC tasks (N=48)**

China: 24 x B1-C candidates recorded responses to 5 tasks

Austria: 24 x B1-C candidates recorded responses to 5 tasks

Eliciting feedback from candidates (in candidates' L1)

Survey (N= 48) → Semi-structured interview (N=12; 6 from China, 6 from Austria – 5 B1, 3 B2, 4 C candidates)

Development of PC and IC tasks

- Identification of **specific elements of PC/IC construct** to be targeted
  (e.g. Nakatsuhara, May, Lam & Galaczi, 2018; Willcox-Ficzere, 2019)

- **Main guiding principles:**
  - importance of **visuals/videos** for context
  - inclusion of **clarification request** (IC tasks)
  - inclusion of 2 'Request' tasks with different **power relationships** (PC tasks)
  - wording of tasks and instructions aligned with **CEFR levels** (EVP/EGP)
    - Candidates: B1, B2, C
    - Target output level = B2 (Input text = B1)

- Development of the test specifications using **the socio-cognitive test specification proforma** (e.g. Weir, 2005; O'Sullivan & Dunlea, 2020)

PC task specifications

Skills focus	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: • 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' • 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 by using language for mitigation • use a wide range of pragmalinguistic devices (e.g. downtoners, hedges, conversational routines) appropriately and in accordance with the level of imposition

Features of the Input / Prompt

Description	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. 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.
Video prompt	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?"

	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)**		

IC task specifications

TEST	Aptis Research Pilot Test	COMPONENT	Speaking	Task	IC (a)/(b)

Features of the Test

Skills focus	The task is designed to tap into four features of IC: *responding to a partner, negotiating towards a joint outcome, interactive listening* and *negotiating meaning*. More specifically, the task measures candidates' ability to: • disagree and put forward a different point of view **effectively** and provide **justification** • effectively **link** their own **contribution** to the **partner's** • work towards a decision by trying to **persuade** the partner • **acknowledge partner's views** • demonstrate they **have been listening** carefully/attentively through responding appropriately to the partner's idea • be able to **clarify/rephrase** their points

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

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

	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 Most lexical items up to B1 according to the English Vocabulary Profile list
Grammatical range	Most grammatical structures up to B1 according to the English Grammar Profile list
Content knowledge	General — Specific
Cultural specificity	Neutral — Specific

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)**

*functions targeted: those highlighted in grey to be finalised after piloting

Data analysis

- **Establishing the CEFR level of candidates:** The benchmarking task (Aptis Task 4) was scored by 2-3 trained raters → **3 levels (B1, B2, C)**

- **Sequential & linguistic analyses:** PC/IC test recordings were transcribed & (after 1-day coding workshop + 2 rounds of reliability checks) analysed for:
  - **interactional moves** *(e.g. acknowledging speakers' point of view, projecting upcoming disagreement, disagreeing using 1st point from the task prompt)*
  - **pragmalinguistic devices** *(e.g. downtoner, upgrader, politeness marker)*
  - → Descriptive stats + qualitatively exemplifying salient features across 3 levels

- **Descriptive stats on survey responses:** across 3 levels & 2 L1 groups

- **Thematic analysis on interview transcripts:** 14 themes identified

PC FINDINGS

PC interactional moves
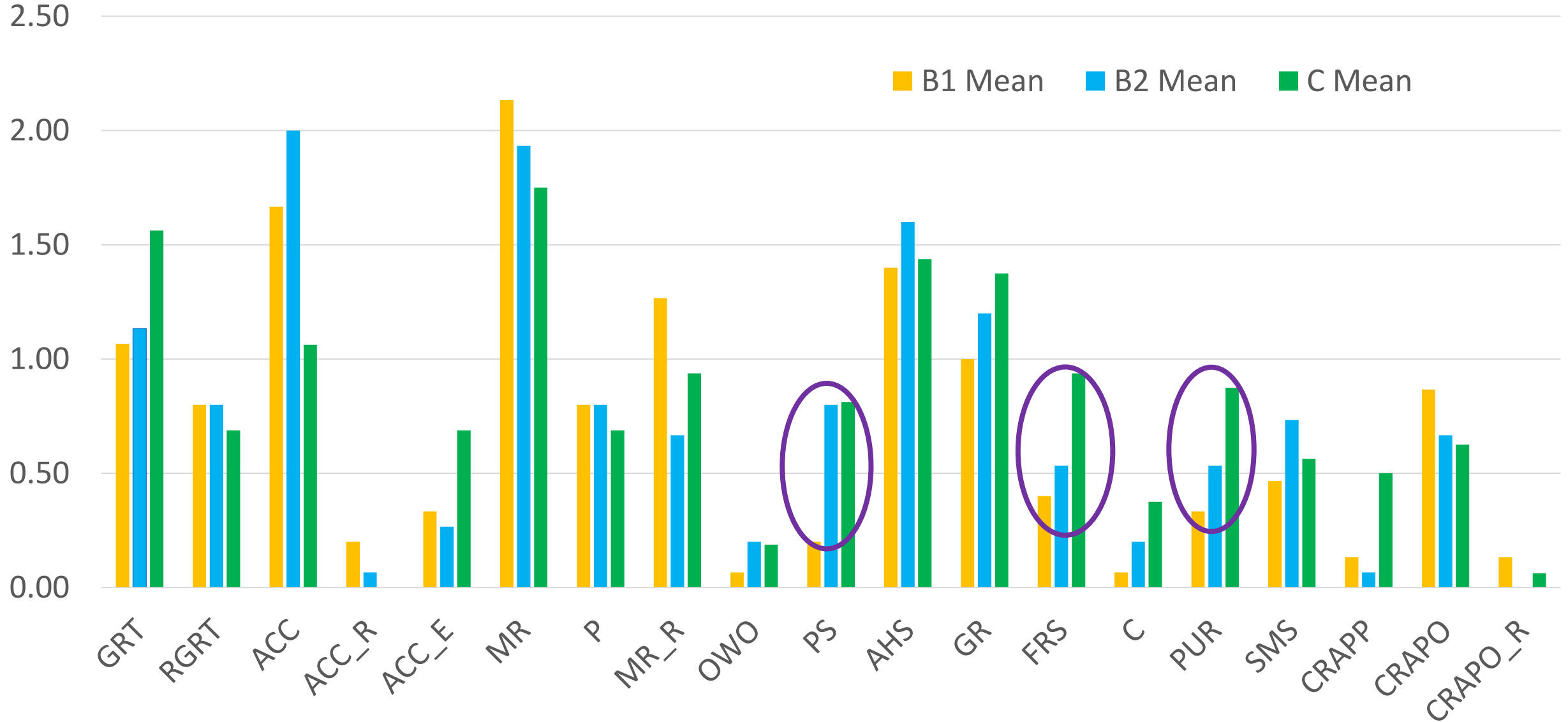
B1 sample (C12)
Task(b): Making a request to a basketball team coach

Hello.	Greeting
I want to, ((clears throat)) ==I want to (.) (erm) back in the team==	Main Request
because (er) (..) I, (erm) my, (..)	Unfinished
(erm) ==maybe you will feel angry about that==, because my (..) (er) faul-, (.) fault. (er)	Acknowledging H's situation
But ==I (.) tried my best to prepare it==. And ==on that day, I had a bad (hand ache).== (erm) I felt so (..) (er) uncomfortable about that. But I, (.) I, but on that situation, ==I can't (er) tell (.) anyone because I (.) didn't want to let the team down==. (erm) And we h-, haven't the (..) some substitute (er) (..) team, teammate. So, (..) (er) so, I, (.) I can't, (.) (erm) so I, (..) (er) I think this is my fault.	Giving an account with elaboration
But I, (.) I f-, (..) I want to...	Unfinished

C sample (A32)

Hi, Mr Swift.	Greeting
So, thank you for taking your time to talk to me.	'Face' related statement
I'm here today to talk about ((clears throat)) (..) (er) last week's game,	Projecting upcoming request
(.) and (.) I (..) first of all, want to thank you for the opportunity to let me play during this (.) first (..) game of mine. And (..) (er) I was really honoured that you put me (.) on the field with the others.	'Face' related statement
(erm) Sadly, I (.) (er) didn't perform as well, and (..) (er) I also think that (erm) you might have had (..) (er) (..) high expectations towards me. But (erm) as it happens, I (er) had a r- (.) really bad headache on that day. And (.) I (.) didn't live up to the expectations and I didn't (erm) like (..) (er) used (er) my whole potential,	Giving an account with elaboration
so I would ask you to (..) (erm) give me another chance,	Main request
even if that may puts you on a tight spot, (.) considering that it may be unfair to the other team players.	Acknowledging H's situation

ALTE – Association of Language Testers in Europe

IC FINDINGS

IC interactional moves

(er) Hi, Jan. (..) (er) (.)	Greeting
I don't agree with you	Explicit disagreement
(.) because you said (.) the (.) group, the groupwork (..) is more likely, is m-, (..) more likely to the real life	Acknowledges Jan's point of view
but	Projects upcoming disagreement- PUD
the group, (..) group (..) presentation only work if all team members work hard.	Disagrees- uses P1
But (in the fact) (..) (er) we (.) can do those things by ourself. (..)	Counter-view
And (er) you say (..) (er) it is a (train) to (..) (train) for the team works skills. (..)	Acknowledges PoV
But (..) have you (..) thought (.) that (.)	PUD
it (..) teamwork maybe a (.) unfair ways to the (..) members, (.) if the, (.) all the members receive the same score?	Disagrees- uses P2
(..) Every members (..) h-, every member (.) do (.) the, (..) do, (..) not do th-, (.) do not (..) (er) have the (..) same (.) contribute to the (.) team.	Elaborates on P2
(er) In fact, (er) many, (..) every members (.) do the, (..) do the different things in a team, (..) such as (..) s-, someone did the, (..) almost everything better. (..) (er) He only (.) received (..) the low score (.) because (.) the, (.) th-, (.) the other one(.) only do (.) an, (.) little things.	Elaborates on last point in response to question

C sample (A31) Task B- travelling solo

Yeah. ==This is a really difficult situation you have there, Dan==. (er) (..) ==I totally understand (.) that (..) you don't really know what to do==.	Acknowledges Dan's situation, appears to concur
So, (.) (er) it is really nice to ==change your plans== spontaneously	Acknowledges Dan's P1
, (er) it is also nice to meet local people and (.) ==yeah, you can probably do that (.) better (.) when you're alone==.	Acknowledges Dan's P2 and appears to concur
However,	Projects upcoming disagreement
(er) ==you can share== (.) those experiences ==with your friends==	Disagrees, using P1
. Just (er) think about (.) ==the fun you can have==,	Elaborates on P1
and (.) ==your friends are supposed to spontaneous as well. So, (er) (.) I think it won't be a problem== (.) ==to change your plans== spontaneously with them	Counter-view
. S-, and, (er) (.) together is always bet-, better. (.) W-, ==so why not?==	Elaborates on P1, Invites change opinion
And also, ==you can== (er) share the costs ==with your friends==,	Disagrees, using P2
which is (.) pretty nice, because Scotland isn't a (.) cheap place to go, ==you know==. And ==with your friends==, (er) (.) it's just (.) less expensive. ==So, I would do it.==	Elaborates on P2
==Yes, of course.==	Agrees to elaborate
(er) (.) ==You know that== (er) Great Britain is not a cheap place. So, (.) (er) probably, it is (.) better to travel ==with your friends== (..) as, (er) for example, a, (er) a hotel costs about a hundred pounds. (erm) For four people, that's just twenty-five and not a hundred. That's cheap.	Elaborates on last point in response to question

Selected survey & interview findings

Selected survey questions	IC tasks
The time I had to prepare for the talk was OK	70.8%
The time I had to speak was OK	58.3%
The time I had to answer Jan and Dan's follow-up Q was OK	66.7%
Instructions for the tasks were clear/very clear	95.9%
I knew who I had to communicate with in these tasks	93.8%
To understand Jan and Dan's points, the video was helpful/very helpful	93.8%
The video helped me to feel that I was communicating with Jan and Dan	54.2%
Trying to persuade someone would happen often/sometimes in real life	89.6%

It was definitely better with videos so if you really see it on the visuals. It was just a bit more human and realistic. (A35, B2)

I think it happens at any time. I think it's quite common in our daily life. (C08, B1)

CONCLUSIONS

University of Bedfordshire

QUT Queensland University of Technology

Main implications

- Findings point to the possibility for semi-direct speaking tasks to elicit:
  - a range of **PC features in requests** (e.g. projecting upcoming request) as well as the extent and structure of moves building up to the main request.
  - a selected range of **IC features**, including acknowledging an interlocutor's view, clarifying/exemplifying a point in order to resolve a breakdown in communication.

- Importantly, candidates 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.

- ➢ **Caveat:** Even within a level there is a range of performances, so the criterial features are **difficult to precisely identify**.

This exploratory study has promising implications for what is possible, in terms of eliciting PC and IC through a semi-direct speaking task.

THANK YOU!

Fumiyo Nakatsuhara & Lyn May