

*Article*

## **Interactional Competence and the Role ‘Role Play’ Plays: The LanguageCert Perspective**

**Leda Lampropoulou**

LanguageCert, UK

Received: 15 September 2021/Accepted: 31 January 2022/Published: 30 March 2022

### **Abstract**

This article describes the importance of including the construct of interactional competence in speaking assessments, drawing mainly from the literature in the field of language testing. The co-construction of meaning and the shared nature of the interaction are seen to be operationalised in an optimal manner using the role play task. The effect of the task is explored through the perspective of the LanguageCert International ESOL Speaking exams, which are used as examples to demonstrate the issues of scalability, discriminability, score separability issues, and the so-called interlocutor effect. Further research and technological innovations will assist in defining and scrutinising the aspects of interactional competence that can be reliably measured.

### **Keywords**

Speaking exams, interaction, oral assessment, role play tasks, LanguageCert

## **1 Introduction**

This article critically summarises the research conducted in the field of interactional competence (IC) in order to describe how the construct of IC has been operationalised in Oral Proficiency Interviews (OPIs) in language testing in LanguageCert examinations. More specifically, the article focuses on describing how the research findings have influenced the format of OPIs, and on explaining the issues and the challenges which have been identified, as well as the issues addressed through the inclusion of role play tasks. The operationalisation of the assessment of IC through the prism of a specific test is considered, through the LanguageCert International ESOL Speaking suite of exams, specifically through the role play task these use as part of the test format.

In Section 2, an historical overview is provided, where the emergence of the concept of interactional competence is briefly outlined. The construct is then defined and its nature is examined, also through its connection with the field of pragmatics. In Section 3, the focus is on oral proficiency interviews and their capacity to tap into the construct of interactional competence. Section 4 looks at the appropriateness of operationalising interactional competence through role play tasks. An example is offered using the

LanguageCert International ESOL Speaking tests. In Section 5, assessment challenges are reviewed, covering the issues of construct scalability, interlocutor effect, score separability, and non-verbal behaviour. A reflection on the inclusion of descriptors pertaining to the assessment of interactional competence features in the LanguageCert International ESOL tests is included. In Section 6, a brief overview of areas that have been identified as key for future development is offered, focusing on the uptake these have started to have. Finally, in Section 7, a conclusion is drawn, summarising the main findings. In the Appendices, the reader will find samples of role play tasks found in the practice papers of the LanguageCert Speaking exams at two different levels, Achiever B1 and Expert C1.

## 2 Background and Definitions

Two decades have passed since Young (2000) described interactional competence (IC) as “a relatively new theory of spoken language use in face-to-face communication” (p.3). It was three decades before then that Hymes (1972) had used the term communicative competence to account for sociocultural variation in language use and acquisition, to challenge Chomsky’s dichotomy between competence and performance, contending that grammar rules cannot exist alone and, therefore lack meaningfulness unless they are considered together with the rules for their functional use. Hymes’ ideas were further developed by Canale and Swain (1980) into an applied linguistics theory which suggested that an individual’s competence includes linguistic competence, discourse competence, pragmatic competence, and strategic competence. L2 teaching practices were tremendously influenced by this theory of communicative competence, and its effect soon extended into language assessment, through Bachman (1990) and Bachman and Palmer (1996), and their observations of the assessment of communicative language ability.

Kramsch (1986) built on Hymes’ theories to develop the construct of what she coined interactional competence, and in defining it she explained that:

*[S]uccessful interaction presupposes not only a shared knowledge of the world, the reference to a common external context of communication, but also the construction of a shared internal context or “sphere of inter-subjectivity” that is built through the collaborative efforts of the interactional partners. (p.367)*

The interpretation of test taker’s speaking performance from this perspective could perhaps alleviate McNamara’s (1997) concern that language assessment based on previous theories considered the test taker’s performance in an unrealistically detached manner, and that the test taker was viewed as the sole liable one for the development of the performance, without considering that they were not the only one participating in it.

The IC construct was also explored by Hall (1995), who focused on interactive practices for which she saw a socially cohesive role for a community, developing through speech acts. Her considerations link pragmatic competence with communicative competence and interactive competence. This link, together with the idea that context is central to the speaking construct, begs the question for a distinction between IC and pragmatics. Young (2011) attempted to answer this by contending that they are interconnected but still distinct competencies. Plough et al. (2018) also identified similarities between IC and pragmatic competence in that they both make use of other competencies, such as grammatical and textual competencies. These are used in parallel as tools to achieve the communication of the intended message, yet IC is emphasised as the skill necessary for “building and maintaining relationships, an aspect of the co-constructed nature of speech” (p.442). The distinction is made even clearer through the understanding that IC emphasises the element of being almost equally constructed by all participants in a discursive practice and is specific to that practice in particular. (Young, 2019).

More importantly, perhaps, Young (2008) notes that IC is not to be found within the individual's skillset or cognitive ability. Young asserts that since participants accomplish the interaction task jointly, the skills described in the theory are distributed among all participants in the interaction. It would, therefore, be inaccurate to claim interactional competence as a skill that a person exercises outside an interaction (He & Young, 1998). Lam (2018), in contrast, asserts that such a skill can only be showcased in the context of a multi-participant interaction which will also rely on the co-participants' performance.

It becomes clear that, to the question featuring in the title of McNamara's (1997) article "*'Interaction' in second language performance assessment: Whose performance?*", the answer can only be, the co-participants' in the interaction.

### 3 Interactional Competence in Language Proficiency Interviews

The assessment of the speaking construct through language proficiency interviews and the extent of operationalisation of IC in different types of oral tests has led to what Galaczi and Taylor (2018) describe as two important strands in theoretical and empirical research, the debates on authenticity and variability.

Roever and Kasper (2018) see a similar 'tug-of-war' between: the conceptualisation of the construct from a primarily psycholinguistic-individualist perspective; and a primarily sociolinguistic-interactional perspective. It is clear that test developers face a dilemma, in which opting for the former perspective focuses on the individual and allows the elicitation of rateable amounts of language samples but can be considered invalid by failing to support inferences on the test taker's ability in typical, real-life interactions. It can, however, be hypothesised that such inferences can be validly supported by speaking tests designed to engage test takers in meaningful, interactive, social situations.

Even before Roever and Kasper's work, language proficiency interviews, such as the Oral Proficiency Interview (OPI), had been castigated for failing to recreate the co-constructed nature of interaction realistically and authentically, and for the absence of salient features of natural conversation caused by the asymmetric relationship between the interlocutors (Young & Milanovic, 1992; Johnson & Tyler, 1998; Johnson, 2001). When speaking tests are based solely on interview-like tasks and conducted in an interview setting, an unequal interaction will occur which will prevent the test from measuring conversational competence in an appropriate manner (Kormos, 1999).

The emerging picture is that if a speaking language test cares to make claims about measuring speaking performances which can be indicative of and generalisable to interactive social contexts, this can only happen through a broadened construct that includes interactional competence operationalised through tasks in which the co-participants jointly engage in conversation. This is the work that LanguageCert are currently developing and trialling.

### 4 The Role 'Role Play' Plays

Paired (and grouped) speaking tests, by nature, allow test takers to interact and co-construct discourse, a strength which, among other reasons too, has made the paired format a common choice not only for classroom-based assessment, but also for high-stakes exams (May, 2011). Moreover, Ockey et al. (2015) suggest that even monologic speaking tests can measure interactional competence with the inclusion of dedicated tasks. Consequently, the onus probandi (burden of proof) appears to fall on task design.

On the one hand, Plough et al. (2018) claim that a unanimous verdict has yet to be reached regarding the extent to which the optimum operationalisation of IC relates to specific speaking task types. On the other hand, the role play task seems to have won the battle between the choice of tasks, as suggested by the findings of several studies.

Kormos (1999) compared non-scripted interviews and guided role play activities in oral assessments using discourse analysis and found that in roleplay “the conversational interaction is more symmetrical” (p.1). Moreover, she established that role play tasks can imitate aspects of conversations in an authentic and realistic manner and found that they can be useful in measuring conversational competence as exhibited in the test takers’ performance, while also concluding that, in terms of measuring conversation management, role play activities can better elicit the manifestation of IC features. Okada’s (2010) findings align with Kormos’ (1999) conclusions. In his study, which discusses roleplay in OPIs in terms of its construct validity, he describes the competencies displayed in performing a role play activity as highly resembling those observed in real-life conversations and he concludes by recognising roleplay as a valid assessment instrument. In a very recent study based on a conversation analysis (CA) of a corpus of role play interaction, Youn (2020) was able to confirm these findings, while maintaining that the language samples elicited through role play interactions, despite not being entirely authentic, can still showcase the test taker’s level of competence regarding how well they would perform in a similar interaction in real life. Hu (2015) also found that roleplay affords an easier access to IC features than other types of paired tasks.

Apart from the conversational characteristics of IC featuring realistically in roleplay, researchers were able to point to more reasons arguing for the inclusion of such tasks in oral proficiency interviews. As an example, in response to the debates on validity and authenticity, Kasper and Youn (2018) assert that roleplay can be used to generate performances with authentic interactional features, such as topic and turn taking management. In addition, attempts to sequence organization attempts, while affording testers the element of control required to make the interaction measurable, render roleplay valid in terms of construct representation. The potential of role play tasks to allow test takers to co-construct discourse is also noted by Galaczi and Taylor (2020).

It is important to note that role play tasks need not be limited to paired speaking test formats, however. OPI roleplays can be conducted with a trained examiner/ interlocutor assuming different roles (Ikeda, 2017; Youn, 2015, 2020). In these OPI roleplays, as in the LanguageCert International ESOL Speaking suite of exams (Appendix 1 & 2), a specific part of the speaking test is dedicated to a role play activity. During that part, the examiner sets the context by informing the test taker of the scenario and the roles to be assumed (as also explained in Kasper & Youn, 2018).

In the case of LanguageCert, the examiner may assume different personas, which range in register formality, such as a colleague or a line manager, a neighbour or a stranger in the street, a doctor’s receptionist or a tour agent, thus enabling different levels of the Common European Framework of Reference for Languages (CEFR) (as analysed by the Council of Europe, 2018) to be measured. Unlike the interlocutor/examiners assuming different roles, the test taker, is *not* expected to take on a role other than their actual self in that interaction. In this context, a high degree of authenticity can be achieved, since the language the test taker will have to use can be expected to resemble the Target Language Use (TLU) domain of social interactions. This is because, in real life, they are likely to need to book a dentist’s appointment or a hotel room, but it would be irrelevant for a non-specific test to assess how well the test takers can perform on the other end of the interaction and assume the role of the doctor or the receptionist, the kind of roles that they may never need to assume in real life.

In the interactions described above, which can either be brief or develop unscripted for a longer period depending on the targeted CEFR level and the test taker’s ability, a wider range of functions can be elicited than the interviewer-structured interaction allows, such as expressing regret, sympathy, condolence, expressing surprise or lack of it, complaining, offering and accepting an apology, etc. (LanguageCert, 2020). The item writer aiming to elicit the demonstration of functional language relating, for instance, to an apology may choose to set the context of the test taker’s late arrival for a meeting with a friend or to work.

Given the evidence above, we can assert with confidence that role play tasks can be considered as appropriately operationalising the construct of interactional competence (Grabowski 2013; Kasper & Youn, 2018; Walters 2007, 2013; Youn 2015).

## 5 Assessing Interactional Competence

Assessment professionals who adopt a sociolinguistic-interactional perspective, foregrounded by the research in applied linguistics, which includes roleplay and other interactive tasks to operationalise the construct of interactional competence, are immediately faced with the challenge of having to assess it. This is a multi-faceted challenge. The reason for the challenge for assessors is that research has revealed two main problematising areas in the measurement of IC. These are, one, the need for differentiation at various levels (or “scalability”); and two, “discriminability” (Galaczi & Taylor, 2018, p. 230), where the separability of scores in the co-constructed performance, also sometimes referred to as “the interlocutor effect” (O’Sullivan, 2002), must deal with the issue of the feasibility of measuring non-verbal behaviour as part of the construct itself.

### 5.1 Descriptors and scalability and discriminability issues

The Common European Framework (CEFR) has developed a descriptive scheme providing scaled descriptors for communicative language competences, which are classified into linguistic competences, sociolinguistic competences, and pragmatic competences (Council of Europe, 2018). The scheme – provided in Appendix 3 – can be read horizontally and vertically, with the horizontal dimension describing the different capabilities expected at the level, while the vertical one attempts to sequence an ascending series of learner proficiency. There is one scale specific to IC, *Interaction*. Table 1 elaborates.

Table 1

*Interaction Scale Across the CEFR Levels*

INTERACTION	
C2	Can interact with ease and skill, picking up and using non-verbal and intonational cues apparently effortlessly. Can interweave his/her contribution into the joint discourse with fully natural turntaking, referencing, allusion making, etc.
C1	Can select a suitable phrase from a readily available range of discourse functions to preface his remarks in order to get or to keep the floor and to relate his/her own contributions skilfully to those of other speakers.
B2	Can initiate discourse, take his/her turn when appropriate and end conversation when he/she needs to, though he/she may not always do this elegantly. Can help the discussion along on familiar ground confirming comprehension, inviting others in, etc.
B1	Can initiate, maintain and close simple face-to-face conversation on topics that are familiar or of personal interest. Can repeat back part of what someone has said to confirm mutual understanding.
A2	Can answer questions and respond to simple statements. Can indicate when he/she is following but is rarely able to understand enough to keep conversations going of his/her own accord.
A1	Can ask and answer questions about personal details. Can interact in a simple way but communication is totally dependent on repetition, rephrasing and repair.

*Note.* Reprinted from <https://www.coe.int/en/web/common-european-framework-reference-languages/table-3-cefr-3.3-common-reference-levels-qualitative-aspects-of-spoken-language-use>

Table 1 above outlines the relevant IC features that learners can be expected to have acquired at each CEFR level, from A1- C2. Oral proficiency tests mapped to the CEFR often use these as a reference tool to describe standard performance expected at each exam level. To use the same example as in the

previous section with the role play task, the LanguageCert International ESOL Speaking test mark scheme describes an A1-A2 test taker as being expected to rely on the support of the interlocutor/examiner. At B1 level, turn taking is expected to be mostly natural, whereas at B2 level, the test taker should be able to handle topic and turn management appropriately and independently, while not always elegantly. These areas are accounted for under Task Fulfilment and Coherence. There are also descriptors under the Pronunciation, Intonation and Fluency criterion/criteria which references the use of intonation to support meaning.

There are, however, no descriptors to cover non-verbal behaviour features, such as eye contact or posture (LanguageCert, 2021). Both have been identified as key features of IC. Such features might be difficult to include in standardised high-stakes exams at present, at least until more research and technological advances permit. However, there are more IC features that could be described and used to measure IC.

The importance an assessment developer places on IC can perhaps be detected by noting whether IC features are displayed under various criteria, informing them by being included in the descriptors, or whether IC is seen as a separate criterion, in a manner that also has a greater impact on the test taker's overall score. In the LanguageCert ESOL Speaking exam the former is the case, but given the prominence that IC is currently being given and the research evidencing its role in communication, it will be interesting to see whether in future revisions of the LanguageCert ESOL Speaking mark schemes IC features will be assigned to a criterion on their own, as has happened with more recently developed LanguageCert exams, e.g., the LanguageCert SELT Speaking & Listening test, in which the criterion is referred to as *Interactive Communication* and *Task Fulfilment*.

The CEFR scales on interaction (Council of Europe, 2018) do include references to some IC features, yet neither consistently nor at all levels. The need for clearer and more specific descriptors differentiating between performance levels has been highlighted across L2 assessment literature relating to IC (Galaczi, 2014; Galaczi & Taylor, 2018, 2020; Lam, 2018; Seedhouse, 2012). These descriptors will need to be developed further, before they can be of wider use to language assessment stakeholders. Furthermore, Galaczi and Taylor (2020), in listing the key features of IC, also refer to breakdown repair, interactive listening, and non-verbal behaviour, aspects analysed in a very limited way in the CEFR scales, even though research relating to rater studies and test taker discourse has noted that they are salient IC features (Ducasse & Brown, 2009; Galaczi, 2014; Gan, 2010; Gan, Davison, & Hamp-Lyons, 2008; May, 2011; Orr, 2002).

Roever and Ikeda (2021) argue that “IC develops along a predictable trajectory” (p.3). They report that research in second language acquisition demonstrates that – as proficiency improves – learners’ IC expands in range and improves in appropriateness (Al-Gahtani & Roever, 2012, 2014, 2018; Cekaite, 2007; Pekarek Doehler, 2019). Still, Roever and Ikeda (2021) identify a challenge in drawing a clear distinction at the higher levels, where IC features may be harder to describe, and this can apply both to L2 and to L1 speakers. This challenge may be illustrated in the descriptors in the mark schemes used at the higher levels of the LanguageCert International ESOL speaking exam where the differentiation – albeit minimal – between the descriptions of turn-taking performance at the two higher CEFR levels exists. At CEFR C1, the criterion for a passing mark at Task Fulfilment and Coherence includes a descriptor of a performance where *turn taking is naturally handled*. Going up a mark at the same criterion, *turn taking needs to be spontaneous, flexible and wholly natural*. Looking at the highest level offered, CEFR C2, the expectation for a passing mark under the same criterion describes turn taking as *naturally handled with a high degree of flexibility* whereas for full marks the descriptor expects turn taking to be *consistently spontaneous, flexible and wholly natural*. It is seen that the differentiation between a passing performance and the one achieving full marks is made through assessing how consistently and flexibly the skill is demonstrated. However, this is not uncommon for mark scheme descriptors aligned to the CEFR, yet it is perhaps indicative of the CEFR’s limitation pertaining to the vague differentiation between IC descriptors

at the different levels, underlined by researchers just above. In addition, in marking examiners' training, the differences in benchmarked performances can be used to standardise what a performance at the level entails, and this is much easier to achieve at a level-specific test, such as the LanguageCert IESOL, than at a multi-level one.

Researchers still maintain that scalability and discriminability are possible. It is understood, as mentioned above, that IC develops in parallel with the learners' general L2 language ability and that as the learners' cognitive processes rely on higher automaticity of conversation processes (Field, 2011), their working memory will afford them a more effective and collaborative participation in interactions (Galaczi, 2014). Roever and Kasper (2018) point to the sequential organization of speech events as a gradable characteristic that can be classified and rated. In their study, they suggest that certain interactional features, such as repair, could be induced by the examiner attempting to elicit this strategy. Galaczi and Taylor (2020) also advise in favour of supporting interaction at the lower CEFR levels with visual or verbal prompts, for reasons of scoring practicality and reliability. Lam (2018) looks at IC through the prism of interactive listening and notes that IC features need to be accounted for as more than the sum of the test taker's responses, and that their appropriacy to the interaction needs to be given prominence.

In discussing role play tasks, Youn (2019, 2020) provides evidence that interactional performances can be elicited so differences can be measurable against rating criteria. This task type also seems to offer itself for appropriately accommodating highly specific professional contexts, such as the context of radiotelephony communication in aviation, where the need to include IC reference to elements of professional knowledge and role behaviour seems to be particularly critical (Kim, 2012). For example, Kim (2012) suggests that the success of the communication in the interaction between pilots and air traffic controllers is so important that the test taker's ability to effectively interact using the aviation radiotelephony conventions should form part of the construct of such an ESP assessment. In such an assessment, the role play would assign the test taker with the role they will be called to operate in in their future jobs, whereas the examiner would take on the persona of the opposite role, to achieve an, as much as possible, authentic performance. The research literature that has been discussed above would appear to indicate that role play tasks are strong contenders for being judged the most effective means by which IC can be measured.

## 5.2 Interlocutor effect and score separability issues

O'Sullivan (2002) used the term 'interlocutor effect' to refer to the sociolinguistics concept of the influence asserted in the interaction by the participants' identities and characteristics. From an assessment perspective, where the focus is traditionally on the individual, the idea and perhaps even the name of *interactional competence* could be enough to raise concern over standardisation and, consequently, validity. At the same time, the co-construction of meaning between the interlocutors perplexes this further, as the test takers' contributions and their performance are seen as shared, interwoven, and linked (Brown, 2003; May, 2011; McNamara, 1997; Roever & Kasper, 2018).

L2 assessment research exploring the different interlocutor variables such as gender, cultural background, acquaintanceship (O'Sullivan, 2002), and extroversion (Nakatsuhara, 2014), did find such characteristics exerting an influence. However, Brown and McNamara (1997) concluded that "the magnitude or direction of that influence is less clear and not directly predictable" (as cited in Galaczi & Taylor, 2020, p.343). More importantly, it is the construct definition that should determine whether this variability is irrelevant and undesired, or whether it is actually part of the construct itself (Galaczi & Taylor, 2020).

Even so, the paired speaking test format can be criticised for (mis)matching test takers of different abilities, causing an observed shared performance that is unrepresentative of the true capabilities of

the individual participants in the interaction. Hu (2015) claims that a more proficient speaker will be disadvantaged if paired with a substantially less proficient speaker.

On the debate on the (in)separability of test takers' scores, May's (2011) suggestion that shared scores could be awarded in response to what raters perceive as a mutually achieved performance has not yet been widely accepted, and although it is a tempting prospect, in high-stakes testing especially, it appears that there is a long way before this can be done, if ever.

For now, the safest path seems to include the challenge of having to overcome the perception of test takers' contributions to the interactions as entangled (Fulcher, 2010) and of training raters to isolate what the individual test taker brings to the paired task. Under this light, raters might be facilitated by a paired task performed between the individual test taker and the interlocutor/examiner, instead of between a pair of test takers. The test takers' contributions can become even more distinguishable and measurable in a role play task, where opportunities for a more symmetrical interaction can be afforded, and the interlocutor/examiner can be trained to elicit specific IC resources the tester is interested in examining.

### 5.3 Non-verbal behaviour

A third issue which needs to be mentioned as a problematising area in assessing IC is non-verbal behaviour, even though it has been considerably less researched in L2 assessment literature, both in terms of its conceptualisation as part of the IC construct, and its operationalisation. Features such as eye contact, facial expression, and posture have been included by Galaczi and Taylor (2020) as denoting non-verbal behaviour pertaining to IC. Researchers have indicated that raters perceive and note non-verbal behaviour even if it is not described in the rating scales (May, 2011; Nakatsuhara et al., 2018; Vo, 2019). Nonetheless, it is still seen as too complex a model to attempt to assess. Oksaar (1990), one of the first explorers of the concept, who was also able to provide insight from multilingual contexts, defined IC aspects with reference to "cultureme and behaviourme" (p.530), which include paralinguistic features as well as sociocultural norms, which, if testers are to include in the construct, they will also need to answer the imminent question: 'whose culture?'

To conclude, integrating IC scales into speaking assessments would appear to enable a wider and more accurate representation of the construct as well as allowing valid inferences about real-world speaking competences, despite the issues which remain under investigation (Roever & Kasper, 2018).

## 6 Developing Research Areas

L2 development of interactional competence (IC) has been widely explored in literature and continues to offer a fertile field for research, while L2 assessment literature has been growing exponentially, and can be expected to continue in a similar manner. The construct of IC is far from having been completely researched, and areas of future research involve both older and newer developments in language testing in general. Plough et al. (2018), see future research targeting four main areas. The first two pertain to issues already touched upon in this literature review, namely the link between task type and elicited evidence of the IC construct, and the role of 'behaviourme'. The other two involve technology-related issues, as in the effect of the mode of speaking test delivery on IC affordances to test-takers, and the extent to which IC inferences can be drawn using computer-delivered tests.

On the first pointer, Youn (2020) argues for the usefulness of CA contributions in recognising various interactional devices in speaking assessment discourse emerging from interactional performances, to inform L2 learning and assessment.

Roever and Kasper's (2018) suggestion incorporates the second and the fourth points, and combines the visual access allowed by computer-assisted testing with using the methodological tools multimodal CA provides, to drive research on non-verbal behaviour such as gaze, gesture, and head movements as part of IC.

On the comparability of the IC construct through different modes of delivery, Nakatsuhara et al. (2017) looked at video-conferencing meetings and how these were distinguished from face-to-face meetings in the use of back channels and the management of turn taking. The researchers noticed differences in the interactions which point to the question of whether IC could support different operationalisations for different delivery modes.

In computer-delivered tests May (2011) also sees the potential for isolating test-takers contributions to co-constructed interactions through a standardised prompt. To these areas, Lam (2018) adds the need for research to support the creation and development of more accurate IC rating scales.

## 7 Conclusion

This article has provided a critical overview of the literature looking at interactional competence as a skill and construct, and its conceptualisation and operationalisation in Oral Proficiency Interviews in language testing using the role play activity as a task type. It has become possible to recognise interactional competence as an important construct pertaining to spoken ability, one that is highly relevant to real-life social contexts. More specifically, IC features such as turn management (e.g., interrupting), interactive listening (e.g., backchanneling), or non-verbal behaviour (e.g., laughter) are seen as key concepts in measuring interactional competence (Galaczi & Taylor, 2020). These have had a varying degree of uptake from assessment developers as some seem easier than others to integrate into assessment tasks, such as turn and topic management. Others, however, seem to require further research or innovative technology before they can be accepted by testers and test stakeholders as measurable and construct-relevant. Non-verbal behaviour or interactive listening are two such areas that require further research.

The literature relating the measurement of interactional competence with a specific task has found the role play activity to be a realistic and authentic task type, able to tap into most of the IC characteristics testers would wish to elicit. However, as there is no overall comparison of all possible tasks, as with a lot of issues in assessment, there are no definite solutions without considering the test purpose and the target language use (TLU) domain. Nevertheless, the role play task has been found to afford a less unequal interaction than other types of tasks, like non-scripted interviews (Kormos, 1999) and through appropriately designed role play situations the power imbalance can be authentically created and simulated, as in a situation between an employee and their manager, or a patient and their doctor.

To better illustrate the roleplay task's effectiveness in assessing IC, the LanguageCert International ESOL test has been used. It appears to be able to operationalise the IC features that assessment developers aim to elicit, in addition to overcoming the challenge of the inseparability of scores, since the performance is shared between the test taker and the interlocutor. The issues identified in the literature that also seem pertinent to the specific assessment of IC relate to the inclusion of more IC features in the mark scheme and the scalability of these, together with the issue of including relevant aspects of non-verbal behaviour.

Looking ahead, more research and empirical studies will allow a stronger integration of IC features in tests measuring speaking constructs. This will most likely be facilitated through technological innovations which will accelerate and enhance assessment design and delivery, as well as allow for a fuller exploration and conceptualisation of interactional competence.

## Appendix 1

### Practice paper 2 of LanguageCert International ESOL (Speaking), B1 level, part 2

LanguageCert Achiever B1

**PART 2 (3 minutes)**

I: Now, Part Two. We are going to role-play some situations. I want you to start or respond.  
First situation (choose one situation from **A**).

**A**

- We're friends. I start.  
*I'm training for a race. Do you want to join me?*
- I work for the local government. I start.  
*Good morning. I'm doing some research. Are you happy with local leisure facilities?*
- I was your teacher last year. I start.  
*Hello, nice to see you again. How's your new class?*
- I'm a stranger in your town. I start.  
*Excuse me, do you know of a good place to eat near here?*

C: (Responds.)

I: (Role-play the situation with the candidate – approximately two turns each.)

I: Second situation (choose one situation from **B**).

**B**

- I work at a library. You want some help. You start.
- We're neighbours. You want me to look after some things while you are away. You start.
- I work at the airline baggage desk. Your bag is missing. You start.
- I work in a shop. I've given you the wrong change. You start.

C: (Initiates.)

I: (Role-play the situation with the candidate – approximately two turns each.)

I: (Role-play a third situation from **A** or **B** if time allows.)

I: Thank you.

Note: Reprinted from <https://www.languagecert.org/en/preparation/practice-material/languagecert-international-esol>

## Appendix 2

### Practice paper 6 of LanguageCert International ESOL (Speaking), C1 level, part 2

LanguageCert Expert C1

**PART 2 (3 minutes)**

I: Now, Part Two. We are going to role-play some situations. I want you to start or respond.  
First situation (choose one situation from **A**).

**A**

- We're flatmates. I start.  
*Why on earth do you keep tidying up? I can't find my laptop yet again.*
- We're classmates. I start.  
*What do you think of our new teacher? She seems a bit strict to me.*
- We're neighbours. I live next door. I start.  
*Your child's football has made a mark on my fence again.*
- You're visiting a museum. I'm the security guard. I start.  
*Excuse me. I've told you twice already – please don't stand so close to the paintings.*

C: (Responds.)

I: (Role-play the situation with the candidate – approximately two turns each.)

I: Second situation (choose one situation from **B**).

**B**

- I'm your boss. You think you deserve a pay raise for your hard work. You start.
- We're flatmates. I haven't paid my share of the rent for the last three months. You start.
- I'm a plumber. I've recently done some work in your house and you're totally dissatisfied. You start.
- We're best friends. I've just told you I've been made redundant. You start.

C: (Initiates.)

I: (Role-play the situation with the candidate – approximately two turns each.)

I: (Role-play a third situation from **A** or **B** if time allows.)

I: Thank you.

Note: Reprinted from <https://www.languagecert.org/en/preparation/practice-material/languagecert-international-esol>

## Appendix 3

### CEFR interaction scale across the CEFR levels

	RANGE	ACCURACY	FLUENCY	INTERACTION	COHERENCE
C2	Shows great flexibility reformulating ideas in differing linguistic forms to convey finer shades of meaning precisely, to give emphasis, to differentiate and to eliminate ambiguity. Also has a good command of idiomatic expressions and colloquialisms	Maintains consistent grammatical control of complex language, even while attention is otherwise engaged (e.g. in forward planning, in monitoring others' reactions).	Can express him/herself spontaneously at length with a natural colloquial flow, avoiding or backtracking around any difficulty so smoothly that the interlocutor is hardly aware of it.	Can interact with ease and skill, picking up and using non-verbal and intonational cues apparently effortlessly. Can interweave his/her contribution into the joint discourse with fully natural turntaking, referencing, allusion making etc.	Can create coherent and cohesive discourse making full and appropriate use of a variety of organisational patterns and a wide range of connectors and other cohesive devices.
C1	Has a good command of a broad range of language allowing him/her to select a formulation to express him/herself clearly in an appropriate style on a wide range of general, academic, professional or leisure topics without having to restrict what he/she wants to say.	Consistently maintains a high degree of grammatical accuracy; errors are rare, difficult to spot and generally corrected when they do occur.	Can express him/herself fluently and spontaneously, almost effortlessly. Only a conceptually difficult subject can hinder a natural, smooth flow of language.	Can select a suitable phrase from a readily available range of discourse functions to preface his remarks in order to get or to keep the floor and to relate his/her own contributions skilfully to those of other speakers.	Can produce clear, smoothly-flowing, well-structured speech, showing controlled use of organisational patterns, connectors and cohesive devices.
B2	Has a sufficient range of language to be able to give clear descriptions, express viewpoints on most general topics, without much con-spicious searching for words, using some complex sentence forms to do so.	Shows a relatively high degree of grammatical control. Does not make errors which cause misunderstanding, and can correct most of his/her mistakes.	Can produce stretches of language with a fairly even tempo; although he/she can be hesitant as he or she searches for patterns and expressions, there are few noticeably long pauses.	Can initiate discourse, take his/her turn when appropriate and end conversation when he / she needs to, though he /she may not always do this elegantly. Can help the discussion along on familiar ground confirming comprehension, inviting others in, etc.	Can use a limited number of cohesive devices to link his/her utterances into clear, coherent discourse, though there may be some "jumpiness" in a long con-tribution.

B1	Has enough language to get by, with sufficient vocabulary to express him/herself with some hesitation and circumlocutions on topics such as family, hobbies and interests, work, travel, and current events.	Uses reasonably accurately a repertoire of frequently used “routines” and patterns associated with more predictable situations.	Can keep going comprehensibly, even though pausing for grammatical and lexical planning and repair is very evident, especially in longer stretches of free production.	Can initiate, maintain and close simple face-to-face conversation on topics that are familiar or of personal interest. Can repeat back part of what someone has said to confirm mutual understanding.	Can link a series of shorter, discrete simple elements into a connected, linear sequence of points.
A2	Uses basic sentence patterns with memorised phrases, groups of a few words and formulae in order to communicate limited information in simple everyday situations.	Uses some simple structures correctly, but still systematically makes basic mistakes.	Can make him/herself understood in very short utterances, even though pauses, false starts and reformulation are very evident.	Can answer questions and respond to simple statements. Can indicate when he/she is following but is rarely able to understand enough to keep conversation going of his/her own accord.	Can link groups of words with simple connectors like “and”, “but” and “because”.
A1	Has a very basic repertoire of words and simple phrases related to personal details and particular concrete situations.	Shows only limited control of a few simple grammatical structures and sentence patterns in a memorised repertoire.	Can manage very short, isolated, mainly pre-packaged utterances, with much pausing to search for expressions, to articulate less familiar words, and to repair communication.	Can ask and answer questions about personal details. Can interact in a simple way but communication is totally dependent on repetition, rephrasing and repair.	Can link words or groups of words with very basic linear connectors like “and” or “then”.

Note. Reprinted from <https://www.coe.int/en/web/common-european-framework-reference-languages/table-3-cefr-3.3-common-reference-levels-qualitative-aspects-of-spoken-language-use>

## References

- Al-Gahtani, S., & Roever, C. (2012). Proficiency and sequential organization of L2 requests. *Applied Linguistics*, 33(1), 42–65. <https://doi.org/10.1093/applin/amr031>
- Al-Gahtani, S., & Roever, C. (2014). Preference structure in L2 Arabic requests. *Intercultural Pragmatics*, 11(4), 619–643. <https://doi.org/10.1515/ip-2014-0027>
- Al-Gahtani, S., & Roever, C. (2018). Proficiency and preference organization in second language refusals. *Journal of Pragmatics*, 129, 140–153. <https://doi.org/10.1016/j.pragma.2018.01.014>
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford University Press.

- Bachman, L. & Palmer, A. (1996). *Language testing in practice*. Oxford University Press.
- Barraja-Rohan, A. (2011). Using conversation analysis in the second language classroom to teach interactional competence. *Language Teaching Research*, 15(4), 479–507 <https://doi.org/10.1177/1362168811412878>
- Brown, A. (2003). Interviewer variation and the co-construction of speaking proficiency. *Language Testing*, 20(1), 1–25. <https://doi.org/10.1191/0265532203lt242oa>
- Brown, A., & McNamara, T. (2004). “The devil is in the detail”: Researching gender issues in language assessment. *TESOL Quarterly*, 38(3), 524–538. <http://doi.org/10.2307/3588353>
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1-47.
- Cekaite, A. (2007). A child’s development of interactional competence in a Swedish L2 classroom. *The Modern Language Journal*, 91(1), 45–62. <https://doi.org/10.1111/j.1540-4781.2007.00509.x>
- Council of Europe (2018). *Common European framework of reference for languages: Learning, teaching, assessment: Companion volume with new descriptors*. Council of Europe.
- Ducasse, A., & Brown, A. (2009). Assessing paired orals: Rater’s orientation to interaction. *Language Testing*, 26(3), 423–443.
- Field, J. (2011). Cognitive validity. In L. Taylor (ed.), *Examining speaking: Research and practice in assessing second language speaking* (pp. 65–111). *Studies in Language Testing* (30). UCLES/Cambridge University Press.
- Fulcher, G. (2010). *Practical language testing*. Routledge.
- Galaczi, E. (2014). Interactional competence across proficiency levels: How do learners manage interaction in paired speaking tests? *Applied Linguistics*, 35(5), 553-574. <https://doi.org/10.1093/applin/amt017>
- Galaczi, E., & 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>
- Galaczi, E., & Taylor, L. (2020). Measuring interactional competence. In Taylor Winke, P., & Brunfaut, T. (Eds.), *The Routledge handbook of second language acquisition and language testing*. Routledge.
- Gan, Z. (2010). Interaction in group oral assessment: A case study of higher- and lower-scoring students. *Language Testing*, 27(4), 585–602.
- Gan, Z., Davison, C., & Hamp-Lyons, L. (2008). Topic negotiation in peer group oral assessment situations: A conversation analytic approach. *Applied Linguistics*, 30(3), 315–334.
- Grabowski, K. (2013). Investigating the construct validity of a role play test designed to measure grammatical and pragmatic knowledge at multiple proficiency levels. In: Ross S.J., Kasper G. (eds), *Assessing second language pragmatics*. Palgrave Advances in Language and Linguistics. Palgrave Macmillan. [https://doi.org/10.1057/9781137003522\\_6](https://doi.org/10.1057/9781137003522_6)
- Hall, J.K. (1995). Aw, man, where you goin? Classroom interaction and the development of L2 interactional competence. *Issues in Applied Linguistics*, 6, 37–62.
- Hall, J. K., & Pekarek Doehler, S. (2011). L2 interactional competence and development. In J. K. Hall, J. Hellermann, & S. Pekarek Doehler (Eds.), *L2 interactional competence and development* (pp. 1-18). *Multilingual Matters*.
- He, A., & Young, R. (1998). Language proficiency interviews: A discourse approach. In R. E. Young & A. He (eds.), *Talking and testing: Discourse approaches to the assessment of oral proficiency* (pp. 1–24).
- Hu, J. (2015). Interaction in assessment-oriented role play: A conversation analytic approach. *Chinese Journal of Applied Linguistics*, 38(4), 472-489. <https://doi.org/10.1515/cjal-2015-0030>
- Hymes, D. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics:*

*Selected readings* (pp. 269-293). Harmondsworth.

- 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 doctoral dissertation]. University of Melbourne, Australia
- Johnson, M. (2001). *The art of nonconversation*. Yale University Press.
- Johnson, M., & Tyler, A. (1998). Re-analyzing the OPI: How much does it look like natural conversation? In R. Young & A. He (eds.), *Talking and testing: Discourse approaches to the assessment of oral proficiency* (pp. 27–51). John Benjamins.
- Kasper, G. (2006). Beyond repair: Conversation analysis as an approach to SLA. *AILA Review*, 19, 83–99.
- Kasper, G. & Youn, S. (2018). Transforming instruction to activity: Roleplay in language assessment. *Applied Linguistics Review*, 9(4), 589-616. <https://doi.org/10.1515/applirev-2017-0020>
- Kim, H. (2012). Exploring the construct of radiotelephony communication: A critique of the ICAO English testing policy from the perspective of Korean aviation experts. *Language Testing and Assessment*, 2(2), 103-110.
- Kormos, J. (1999). Simulating conversations in oral-proficiency assessment: A conversation analysis of role plays and non-scripted interviews in language exams. *Language Testing*, 16(2), 163–188. <https://doi.10.1177/026553229901600203>
- Kramersch, C. (1986). From language proficiency to interactional competence. *The Modern Language Journal*, 70(4), 366-372. <https://doi.10.2307/326815>
- 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.10.1177/0265532218758126>
- LanguageCert, (2020). *LanguageCert international ESOL qualification handbook (speaking)*. <https://www.languagecert.org/en/preparation/practice-material/languagecert-international-esol>
- LanguageCert, (2021). *Assessing speaking performance*. <https://www.languagecert.org/en/preparation/practice-material/languagecert-international-esol>
- Markee, N. (2019). Some theoretical reflections on the construct of interactional competence. <https://doi.10.4324/9781315177021-3>
- May, M. (2011). Interactional Competence in a paired speaking test: Features Salient to raters. *Language Assessment Quarterly*, 8(2), 127-145. <https://doi.10.1080/15434303.2011.565845>
- McNamara, T. F. (1997). ‘Interaction’ in second language performance assessment: Whose performance? *Applied Linguistics*, 18(4), 446-466.
- Nakatsuhara, F. (2014). The co-construction of conversation in group oral tests. Peter Lang.
- Nakatsuhara, F., Inoue, C., Berry, V., & Galaczi, E. (2017). Exploring the use of video-conferencing technology in the assessment of spoken language: A mixed-methods study. *Language Assessment Quarterly*, 14, 1 - 18.
- Nakatsuhara, F., May, L., Lam, D., & Galaczi, E. D. (2018). Learning oriented feedback in the development and assessment of interactional competence. *Research Notes*, 70. <https://www.cambridgeenglish.org/Images/517543-research-notes-70.pdf>
- Ockey, G. J., Koyama, D., Setoguchi, E., & Sun, A. (2015). The extent to which TOEFL iBT speaking scores are associated with performance on oral language tasks and oral ability components for Japanese university students. *Language Testing*, 32(1), 39–62. <https://doi.org/10.1177/0265532214538014>
- Okada, Y. (2010). Role play in oral proficiency interviews: Interactive footing and interactional competencies. *Journal of Pragmatics*, 42, 1647–1668.

- Oksaar, E. (1990). Language contact and culture contact: Towards an integrative approach in second language acquisition. In H. Dechert (Ed.), *Current trends in European second language acquisition research* (pp. 230–43). Multilingual Matters.
- Orr, M. (2002). The FCE Speaking test: Using rater reports to help interpret test scores. *System*, 30(2), 143–154.
- O’Sullivan, B. (2002). Learner acquaintanceship and oral proficiency test pair-task performance. *Language Testing*, 19(3), 277–295. <https://doi.org/10.1191/0265532202lt205oa>
- Pekarek Doehler, S. (2019). On the nature and the development of L2 interactional competence: State of the art and implications for praxis. In M. R. Salaberry & S. Kunitz (Eds.), *Teaching and testing L2 interactional competence: Bridging theory and practice* (25–59). Routledge.
- Plough, I. (2018). Revisiting the speaking construct: The question of interactional competence. *Language Testing*, 35(3), 325–329. <https://doi.org/10.1177/0265532218772322>
- Plough, I., Banerjee, J., & Iwashita, N. (2018). Interactional competence: Genie out of the bottle. *Language Testing*, 35(3), 427–445. <https://doi.org/10.1177/0265532218772325>
- Roever, C., & Ikeda, N. (2021). What scores from monologic speaking tests can(not) tell us about interactional competence. *Language Testing*. <https://doi.org/10.1177/02655322211003332>
- 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/10.1177/0265532218758128>
- Seedhouse, P. (2012). What kind of interaction receives high and low ratings in oral proficiency interviews? *English Profile Journal*, 3, 1–24. <https://doi.org/10.1017/S2041536212000025>
- Van Moere, A. (2006). Validity evidence in a university group oral test. *Language Testing*, 23(4), 411–440. <http://doi.org/10.1191/0265532206lt336oa>
- Vo, S. T. (2019). *Effects of task types on interactional competence in oral communication assessment*. (Unpublished doctoral dissertation). Iowa State University, United States.
- Walters, S. F. (2007). A conversation-analytic hermeneutic rating protocol to assess L2 oral pragmatic competence. *Language Testing*, 27(2), 155–183.
- Walters, S. F. (2013). Interfaces between a discourse completion test and a conversation analysis informed test of L2 pragmatic competence. In Steven J. Ross & Gabriele Kasper (eds.), *Assessing second language pragmatics* (pp. 172–195). Palgrave Macmillan.
- Youn, S. J. (2015). Validity argument for assessing L2 pragmatics in interaction using mixed methods. *Language Testing*, 32, 199–225.
- Youn, S. J. (2019). Interactional features of L2 pragmatic interaction in role play speaking assessment.
- Youn, S. J. (2020). Managing proposal sequences in role play assessments.
- Young, R. F. (2000). Interactional competence: Challenges for validity. Paper presented at the annual meeting of the American Association of Applied Linguistics, Vancouver, Canada.
- Young, R. F. (2008). *Language and interaction*. Routledge.
- Young, R. F. (2011). Interactional competence in language learning, teaching, and testing. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 426–443). Routledge.
- Young, R. F. (2019). Interactional competence and L2 pragmatics. In R. F. Young (ed.), *The Routledge handbook of second language acquisition and pragmatics* (pp. 93–110). Routledge.
- Young, R., & Milanovic, M. (1992). Discourse variation in oral proficiency interviews. *Studies in Second Language Acquisition*, 14, 403–424. <https://doi.org/10.1017/S0272263100011207>

**Leda Lampropoulou** is Head of Assessment at LanguageCert, where she coordinates the development of speaking assessments. She holds a degree in English with Philosophy from the University of London, she is CELTA qualified and has worked in the field of EFL from a range of assessment development positions. She is currently engaged in researching online oral assessment, also as part of her postgraduate degree in Language Testing from Lancaster University. ORCID: 0000-0002-6977-400X