© 2026 International Conference on Multidisciplinary Sciences and Innovations . All Rights Reserved.

GIVING INSTRUCTIONS IN L2 SPEAKING ASSESSMENT: THE ROLE OF INSTRUCTIONAL DISCOURSE IN ELICITING SPEAKING PERFORMANCE

Mamasoliyeva Sarvinoz Hasanjon qizi
Independent reasearcher ASIFL
DOI:https://doi.org/10.5281/zenodo.23161377

Abstract

Instructional discourse — the language examiners and test materials use to direct test-takers through a speaking assessment task — has received comparatively little theoretical attention relative to task design and rating criteria. Yet the way instructions are delivered, whether scripted verbatim or improvised, direct or mitigated, may shape the very interaction from which speaking proficiency is inferred. This paper synthesizes research on the interlocutor frame, examiner variation, and discourse-based validation in L2 speaking assessment to argue that instruction-giving should be treated as a pragmatic and discourse-analytic object of study in its own right. Drawing on the constructs of interactional co-construction, examiner deviation, and discourse-based task validation, the paper proposes a conceptual framework distinguishing three dimensions of instructional discourse — directness, elaboration, and consistency — and discusses their implications for test fairness, examiner training, and future empirical research.

Keywords: instructional discourse, L2 speaking assessment, interlocutor frame, pragmatics, language testing

Introduction

Second language (L2) speaking assessment depends on an interactional event: an examiner elicits performance from a test-taker through a sequence of instructions, prompts, and follow-up questions. Decades of validation research have concentrated on two components of this event — the design of tasks and the criteria used to rate the resulting performance — while the discourse used to introduce and manage the task itself has been comparatively neglected as an object of theoretical inquiry. This is a notable omission, because the instruction is not a neutral delivery mechanism standing outside the assessed interaction; it is part of the interaction, and its form can shape what the test-taker subsequently produces.

The clearest institutional acknowledgment of this problem is the interlocutor frame (IF), a scripted protocol introduced by Cambridge ESOL to standardize what examiners say across test administrations, so that all test-takers participate in a comparable event. The very existence of the IF reflects an implicit theory: that variation in instructional wording is consequential enough to require control. Empirical work has since shown that examiners do deviate from the IF under live testing conditions, and that such deviations can be classified and traced to their effects on candidate output. This body of work establishes that instructional discourse is empirically tractable, but it has developed largely as a quality-assurance concern internal to testing boards rather than as a theorized object within pragmatics or discourse analysis.[2]

This paper takes a conceptual approach: rather than reporting new data, it synthesizes existing discourse-analytic and language-testing literature to argue for treating instructional discourse in L2 speaking assessment as a distinct pragmatic phenomenon, and to propose a framework for describing its variation.

© 2026 International Conference on Multidisciplinary Sciences and Innovations . All Rights Reserved.

2. Theoretical Background

The interlocutor frame was designed to ensure that all test events conform to a common test design, so that test-takers participate in essentially the same event regardless of which examiner conducts the interview (Cambridge ESOL). Despite this scripting, research has repeatedly documented that examiners deviate from the frame in practice. Early work identified and classified such deviations, while a subsequent large-scale study of the IELTS Speaking Test analyzed sixty recorded test events, identifying specific sources of deviation from the IF and comparing candidates' pre- and post-deviation output in terms of elaboration, accuracy, complexity, and fluency. Notably, that study found that examiner deviations from the interlocutor frame had little significant impact on candidate language overall, a finding that itself invites theoretical explanation: it suggests either that candidates are robust to instructional variation, or that the deviations studied were too narrow in scope to capture the pragmatic features most likely to matter.[4]

A complementary line of research reframes the oral proficiency interview itself as co-constructed discourse rather than a one-directional elicitation device. Discourse in dyadic native–non-native speaker interaction has been modeled along dimensions such as interactional contingency, goal orientation, and dominance, with findings showing that examiners tend toward greater goal-orientation while candidates respond more reactively — and that this structure varies systematically with the examiner, the task, and the interview theme. In a related vein, interviewer variation has been shown to be part of how speaking proficiency is co-constructed in the interaction, rather than a source of noise external to it. Taken together, this literature suggests that instruction-giving is not merely a trigger for a separately occurring performance, but a constitutive part of the discourse event being rated.[1]

Discourse-Based Validation of Speaking Tasks

A separate but related tradition uses discourse analysis, including conversation analysis, to validate speaking task types and enrich rating criteria, examining discourse between test-takers in paired or group tasks and between test-takers and examiners in interview-based formats. This tradition has addressed how discourse analysis can inform test construct definition, task validation, and the scrutiny of rater discussions, and has been extended to integrated listening-speaking tasks, where discourse-based analysis of test-taker output has been used to investigate task validity.[3] This literature demonstrates the methodological feasibility of applying discourse-analytic tools to assessment talk, but it has focused primarily on test-taker output and rater behavior rather than on the pragmatic structure of examiner instruction-giving as its own analytic target.

Toward a Conceptual Framework

Synthesizing these three strands, this paper proposes that instructional discourse in L2 speaking assessment can be conceptually organized along three dimensions:

Directness. The degree to which an instruction is delivered as an explicit directive versus a mitigated or indirect request, drawing on speech act and politeness theory. Scripted interlocutor frames typically standardize directness, yet examiners' live deviations may shift an instruction along this dimension without altering its propositional content.[5]

Elaboration. The extent to which an examiner supplements a scripted prompt with additional explanation, examples, gesture, or back-up questions for reticent candidates — a practice explicitly built into some interlocutor frames but left to examiner discretion in its precise extent and timing.[6]

Consistency. The degree to which instructional discourse is delivered uniformly across examiners, candidates, and administrations. This dimension is the direct concern of interlocutor-frame research and examiner-deviation studies, and connects instructional discourse most directly to test fairness.[7]

© 2026 International Conference on Multidisciplinary Sciences and Innovations . All Rights Reserved.

This framework reframes existing findings — the limited measured impact of IF deviation on candidate language, the co-constructed nature of interview discourse, and the discourse-analytic validation of task types — as evidence bearing on specific, separable dimensions of instructional discourse, rather than as a single undifferentiated construct of "examiner variation."

Discussion and Implications

Treating instructional discourse as a pragmatic object with distinct dimensions carries several implications. First, for examiner training, it suggests that standardization efforts might productively target directness and elaboration separately, since a script can hold propositional content constant while still permitting variation in mitigation or additional explanation. Second, for test fairness, the consistency dimension connects most directly to validity arguments: if instructional discourse varies systematically with examiner or candidate characteristics, this variation becomes a source of construct-irrelevant variance requiring monitoring, in the same spirit as existing interlocutor-frame quality assurance. Third, for methodology, the discourse-analytic and conversation-analytic tools already used to validate task types and enrich rating criteria could be extended upward, from test-taker output to the instruction-giving turns that precede it, treating the instruction itself as transcribable, codeable data.

The framework also highlights a gap in the reviewed literature: existing empirical work on examiner deviation has tended to measure downstream effects on candidate language as a whole, rather than isolating which dimension of instructional discourse — directness, elaboration, or consistency — is responsible for an observed effect, or its absence. Future empirical work could test the three-dimensional framework directly, for example by systematically varying directness or elaboration while holding task content constant, and by extending such comparisons across languages and testing contexts beyond the Anglophone, internationally administered examinations that dominate the existing literature.

Conclusion

Instructional discourse in L2 speaking assessment sits at the intersection of pragmatics, discourse analysis, and language testing, yet it has been studied mainly as a quality-assurance problem rather than a theorized pragmatic phenomenon. By synthesizing research on the interlocutor frame, examiner variation, co-constructed interview discourse, and discourse-based task validation, this paper has proposed a three-dimensional conceptual framework — directness, elaboration, and consistency — for describing how instructions are given in speaking assessment and why that variation matters for validity and fairness. Situating instruction-giving within this framework offers a basis for future empirical research that treats the instruction itself, and not only the performance it elicits, as a legitimate object of discourse-analytic inquiry.

References

1. Brown, A. (2003). Interviewer variation and the co-construction of speaking proficiency. Language Testing, 20(1), 1–25.
2. Cambridge English. (n.d.). Interlocutor frames and speaking test administration guidance. Cambridge Assessment English.
3. Frost, K., Elder, C., & Wigglesworth, G. (2011). Investigating the validity of an integrated listening-speaking task: A discourse-based analysis of test-takers' oral performances. Language Testing, 29(3), P 369.
4. IELTS Partners. (n.d.). The impact on candidate language of examiner deviation from a set interlocutor frame in the IELTS Speaking Test [Research report]. IELTS Research Reports.

© 2026 International Conference on Multidisciplinary Sciences and Innovations . All Rights Reserved.

5. Lam, D. (2025). Discourse analysis in language assessment. In The Wiley Handbook of Applied Linguistics (Major Reference Works). Wiley Online Library.
6. Lazaraton, A. (1996). Interlocutor support in oral proficiency interviews: The case of CASE. Language Testing, 13(2), P172.
7. Ross, S., & Berwick, R. (1992). The discourse of accommodation in oral proficiency interviews. Studies in Second Language Acquisition, 14(2), P176.