

Assessing speaking ability in academic context: Focusing on a mixed methods approach

Simin, Shahla ✉

University of Isfahan, Iran (shahlasimin@yahoo.com)

Tavakoli, Mansoor

University of Isfahan, Iran (mr.tavakoli14@gmail.com)



ISSN: 2243-7754
Online ISSN: 2243-7762

OPEN ACCESS

Received: 12 March 2014

Revised: 29 April 2014

Accepted: 1 May 2014

Available Online: 30 May 2014

DOI: 10.5861/ijrsl.2014.758

Abstract

The ability to speak in a foreign language is at the very heart of what it means to be able to use a foreign language. Speaking skill is an important part of the curriculum in language teaching, and this makes it an important object of assessment as well. This study consists of two phases, in the first phase, the qualitative phase, the purpose is to examine whether EFL learners use features of real-world listening input such as time creation device (pause fillers), facilitation device (fixed and conventional phrases), stock phrases and compensation device such as redundancy (repetition, reformulation, and rephrasing) in their natural speech production and in the second phase, the quantitative one, the purpose is to ascertain whether there is any correlation between the participants' use of these features with their proficiency. For this purpose, thirty four EFL students, who were third-year majors in English literature at the University of Isfahan, participated in the interview for their oral production final exam. According to their scores they were divided into two groups, high proficient and low proficient. After the corpus analysis, the results show that time creation device and then stock phrases are the two most frequently used strategies among which time creation fillers are occupying the absolute majority of all strategies employed. The results also reveal that there is no statistically significant difference between two groups regarding their proficiency levels and frequency of appropriate use of spoken features. The immediate implications that can be drawn from the results obtained in this study is that it sheds light on using oral assessment as a necessary and practical way to enhance EFL learners' speaking skills and ability.

Keywords: oral proficiency; oral language assessment; oral language instruction; EFL speaking skills

Assessing speaking ability in academic context: Focusing on a mixed methods approach

1. Introduction

The ability to speak in a foreign language, as noted by Luoma (2004), is at the very heart of what it means to be able to use a foreign language. Our personality, our self-image, our knowledge of the world and our ability to reason and express our thoughts are all reflected in our spoken performance in a foreign language. She also stated that

speaking is also the most difficult language skill to assess reliably. A person's speaking ability is usually judge during a face-to-face interaction, in real time, between an interlocutor and a candidate. The assessor has to make instantaneous judgments about a range of aspects of what is being said, as it is being said. This means that the assessment might depend not only upon which particular features of speech (e.g. pronunciation, accuracy, fluency) the interlocutor pays attention to at any point in time, but upon a host of other factors such as the language level, gender, and status of the interlocutor and the personal characteristics of the interlocutor and candidate. Moreover the nature of the interaction, the sorts of tasks that are presented to the candidate, the questions asked, all will have an impact on the candidate's performance. Therefore, assessing speaking is not impossible but it is difficult (p. ix).

Speaking in a second language has been considered the most challenging of the four skills given the fact that it involves a complex process of constructing meaning (Celce-Murcia & Olshtain, 2000). It is also the most complex skill because of simultaneous monitoring and planning of utterances. This process requires speakers to make decisions about why, how and when to communicate depending on the cultural and social context in which the speaking act occurs (Burns & Seidlhofer, 2002).

In addition, what can be understood is a function of the listener's background and ability as well as those of the speaker. Another difficulty is separating the listening skill from the speaking skill. In spite of the difficulties in testing speaking, it can be very beneficial in that it encourages the teaching of speaking in class. Reading aloud, conversational exchanges, and tests using visual material as stimuli are common test items for testing speaking. Oral interviews, role play tests, and group or pair activities are also useful. One of the great difficulties in testing speaking is the assessment and it's scoring. If possible, the speaking tasks should be recorded and the scoring done from the tape. Aspects of speaking that might be considered in the assessment scale are grammar, pronunciation, fluency, content, organization, and vocabulary.

2. Background

One of the major objectives of teaching oral communication is enhancing students' ability to use oral language in various sociolinguistic contexts. Speaking is often interactive, involving more than one person at the same time. Of course, speaking can be monologic, involving only one speaker, as in a lecture or a radio broadcast. Even if we limit our contexts to academic settings, sociolinguistic conditions and levels of student performance vary significantly. Some students seem to be good at monologue type speeches or oral presentations in tests. Others appear to be more skillful in handling dialogue type interview tests. Still others perform best in discussion activities (Brown 2003; Bonk & Ockey, 2003).

2.1 *Speaking as meaningful interaction*

Speaking and spoken interaction- Teaching and testing experts often talk about speaking as a technical term to refer to one of the various skills that language learners should develop and have. According to Luoma,

this type of speaking tends to be seen as something that individuals do. It is legitimate, and for educational purposes useful, to see speaking in this way too, because it is true that individuals speak, and an important part of language use is personal. Nevertheless, it is also important to remember that speaking forms a part of the shared social activity of talking (Luoma, 2004, p. 20).

2.2 Progression of speaking ability

Fluency - Fluency is a thorny issue in assessing speaking. This is partly because the word 'fluency' has a general meaning, as in 'she is fluent in five languages', and a technical meaning when applied linguists use it to characterize a learner's speech. However, even in technical terminology, fluency can be used in a range of senses. The narrowest definitions only include a few features, typically pausing, hesitations and speech rate, whereas the broadest uses are virtually synonymous with 'speaking proficiency'. Unless the term is defined explicitly, it is simply not clear what a speaker or writer means by it (Freed, 1995; Fulcher, 1996). Definitions of fluency often include references to flow or smoothness, rate of speech, absence of excessive pausing, absence of disturbing hesitation markers, length of utterances, and connectedness (Koponen, 1995). These characterizations are complex, however, because they are not simply descriptions of a speaker's speech but also of a listener's perception of it.

One central part of fluency is related to temporal aspects of speech, such as speaking rate, speech-pause relationships, and frequency of dysfluency markers such as hesitations, repetitions and self-corrections. These can be evaluated by machine and by human impression. Both kinds of studies indicate that when speakers become more fluent their speech rate increases and the speech flow contains fewer pauses and hesitations (Lennon, 1990; Freed, 1995).

They also pause at semantically sensible places, which listeners perceive as the speakers' planning the content of what they are saying rather than groping for words. More fluent speakers tend to speak more and their phrases are longer (Luoma, 2004, p. 89).

In addition to time-bound speed and pausing phenomena, fluency is related to the way that speakers use words, and in particular 'smallwords' such as *really*, *I mean*, and *oh* (Hasselgren, 1998). To focus on the more lexical aspects of fluency, Hasselgren defined it as:

the ability to contribute to what a listener, proficient in the language, would normally perceive as coherent speech, which can be understood without undue strain, and is carried out at a comfortable pace, not being disjointed or disrupted by excessive hesitation (Hasselgren, 1998, p. 155).

She suggested that smallwords are significant in this because they help speakers produce relevant turns and understand the relevance of other speakers' contributions. She summarizes the tasks of smallwords in the following list:

1. They express the communicative intention of the speaker, with respect to what is to be communicated and how it affects the interactional roles of the participants.
2. They point to the textual context in which an utterance has relevance.
3. They indicate the cognitive effect of the previous utterance
4. They enrich the explicature of an utterance, notably by indicating degree of commitment and vagueness.
5. They indicate the state of success of the communication, acknowledging this, or appealing for confirmation, or assistance in bringing it about.

Pragmatic skills - Since the rise of the communicative approach to language learning in the 1970s, learners' ability to *use* a language has been a central focus of attention. However, many of the early studies of SLA focused on learner's acquisition of syntax rather than pragmatics (Kasper, 1996). When studies on inter-language pragmatics began to appear in the late 1980s and 1990s, most of them focused on differences between learners and native speakers rather than the development of pragmatic skills. Recently, initiatives by Kasper (1996), House (1996) and Bardovi-Harlig (1999) have begun to change this, and the early findings suggest that it is possible to distinguish and describe varying levels of control of pragmatic skills. Rather than using scales for something called 'pragmatic skills', it is more common for speaking assessments to focus on areas of it. The *Common European Framework*, for example, considers pragmatic skills to be composed of discourse and functional competence. It proposes six illustrative scales for this, entitled flexibility to circumstances, turn taking, thematic development (in a narrative), coherence and cohesion, fluency, and propositional precision (Council of Europe, 2001, p. 123-129).

Learner grammar - Almost all speaking criteria make some reference to grammar, either as a part of holistic descriptors or in a separate analytic rating criterion. For example, all the ACTFL descriptors mentioned the grammatical features of the learner's speech at different levels, and the CEF analytic grid included a separate criterion called 'Accuracy'. Research also indicates that raters tend to pay a lot of attention to grammar even if the test uses several analytic criteria (Brown, 2000; McNamara, 1996).

Much of the work on common learning orders in SLA has focused mainly on grammar. While there have been arguments in the past about

whether common learning orders can be explained by nature or nurture, and whether early language learning is qualitatively different from the learning of languages later in life, most recent theories propose that people's language learning is usage-based. In other words, all language learning can ultimately be explained by the frequency with which we hear and see different language-use patterns (Luoma, 2004, p. 93).

Ellis(2002), for instance, summarizes studies from a wide range of areas of areas including phonology, reading, spelling, lexis, morphosyntax and formulaic language, which show that the more exposure learners have had to patterns in these areas, the better they know them. This approach proposes that the processes governing language learning are universal, but it also suggests that there is inherent variation in individual language users' proficiencies due to their different experiences of language use. They come from different language backgrounds, they may or may not know other languages, and they are familiar with different ranges of language-use situation. In this vein, Brown (1996, p.188) summarizes the usage-based view of language knowledge as “the constantly updated memory of all linguistic input and output that a particular speaker has ever produced or understood, a memory whose interconnections and salience are constantly being changed and updated as new linguistic experiences are encountered”.

2.3 Spoken words

Considering the rating scales, many rating scales for speaking, according to Luoma, include descriptions of vocabulary use, and at the highest levels these often talk about being able to express oneself precisely and providing evidence of the richness of one's lexicon. This can indeed be important in professional contexts or when trying to convey detailed information. Well-chosen phrases can also make descriptions or stories vivid, and learners who can evoke the listener's feelings deserve to be credited for their ability. However, very 'simple' and 'ordinary' words are also very common in normal spoken discourse, and using these naturally in speech is likewise a maker of highly advanced speaking skills. Moreover, there is a core of phrases and expressions that are highly typical for speaking, which contribute to the listener's impression of the speaker's fluency. They work at the interpersonal level by keeping the conversation going and developing

the relationship between the speakers. This aspect of word use should also be rewarded in assessing speaking (Luoma, 2004, p. 16).

Fixed phrases, fillers and hesitation markers - Speakers need to know words, phrases, and strategies for creating time to speak. According to Luoma,

these are sometimes called fillers or hesitation markers, and they include expressions such as ah, you see, kind of, sort of, and you know, as well as whole expressions such as That's a good question, or Now let me see. Speakers often also use repetition of their own words, or of those used by the previous speaker, to achieve the same purpose, i.e. to keep the floor while formulating what they want to say. These expressions are very common in native speaker speech, but for some reason their appearance in test performances by foreign language learners is sometimes frowned upon (Luoma, 2004, p. 18).

Fixed conventional phrases are also used for other purposes in talk than creating time. Examples of these include responses like *I thought you'd never ask* or *I'm doing all right, all things considered*. The phrases either always have the same form, or they constitute a formula where one or two slots can be filled by various terms (e.g. *what a nice thing to say, what a horrible thing to say*). They have been called lexicalized sentence stems by Pawley and Syder (1983), and lexical phrases by Nattinger and DeCarrico (1992). They are easy for speakers to use because they come almost automatically when a relevant situation arises and because, once a speaker begins such a phrase, saying it will give them time to judge the situation, perhaps plan how they want to put what they want to say next, or think of something else to say.

2.4 Speaking assessment

Testing oral proficiency has become one of the most important issues in language testing since the role of speaking ability has become more central in language teaching (Hartley & Sporing, 1999). However, testing speaking is difficult and cannot be assessed as precisely and easily as other language skills (Harris, 1969). It takes considerable time, effort and training (Hughes, 2003). Farhady, Jafarpour, and Birjandi (1994) however, assert that despite the difficulty of evaluating speaking tests, they should be designed, administered and evaluated regardless of how much cost or effort might be needed.

3. Purpose and research questions

The Purpose of this study is twofold. First, it aimed to investigate whether EFL learners use features of real-world listening input such as time creation device (pause fillers), facilitation device (fixed and conventional phrases), stock phrases and compensation device such as redundancy (repetition, reformulation, and rephrasing) in their natural speech production, if yes, to what extent they use such devices appropriately and whether it is correlated with their language proficiency.

In the view of the facts stated above, the research questions addressed in this study are as follows:

- Do EFL learners use features of real-world listening input such as time creation device (pause fillers), facilitation device (fixed and conventional phrases), stock phrases and compensation device such as redundancy (repetition, reformulation, and rephrasing) in their natural speech production?
- Is there any correlation between the participants' use of spoken features with their language proficiency?

4. Method

4.1 Participants

Thirty four EFL students, both male and female, aged between nineteen to twenty-nine years old, who were third-year majors in English literature at the University of Isfahan, Iran participated in the interview for their oral production final exam. They were highly motivated for the interview test since the score was for their finals. And also the most obvious approach to oral testing and the one presumed to be most valid is the oral interview. An interview provides a very direct method of challenging someone to speak; and it offers a realistic situation in which to assess overall oral mastery of a particular language (Oller & Perkins 1977).

4.2 Materials and procedures

This study was conducted using both qualitative and quantitative method with thirty-four EFL students, all participated in an interview with the researcher. In order to assess the participants' general speaking ability, a version of IELTS interview tests (2007) was utilized at the end of the term as the students' oral production final exam. The participants were assessed on their use of spoken English to answer short questions, speak at length on a familiar topic and also to ask questions and interact with the examiner. All of the procedures suggested by IELTS were observed, such as introduction, extended discourse, elicitation, speculation and attitudes and conclusion. IELTS results were scored on a nine-band scale. This band consists of different levels from zero to native like ability to communicate in spoken English.

Their interviews were tape-recorded and then transcribed. According to their level of proficiency, subjects were divided into two groups: high proficient and low proficient. The purpose was to examine whether these EFL learners use features of real-world listening input such as time creation device (pause fillers), facilitation device (fixed and conventional phrases), stock phrases and compensation device such as redundancy (repetition, reformulation, and rephrasing) in their natural speech production, if yes, to what extent they use such devices appropriately and whether it is correlated with their language proficiency.

During the corpus analysis, features of real-world listening input such as time creation device (pause fillers), facilitation device (fixed and conversational phrases), stock phrases and compensation device such as redundancy (repetition, reformulation, and rephrasing) in the speech production of students were marked and the number of them was counted. The frequencies obtained were described in terms of percentages and frequencies (see Table 2).

5. Results and discussion

As mentioned above, the purpose of the study was to investigate whether EFL learners use features of real-world listening input such as time creation device facilitation device, stock phrases and compensation device in their natural speech production, if yes, to what extent they use such devices appropriately and whether it is correlated with their language proficiency. In order to answer the aforementioned research questions, a mixed method approach, using both qualitative and quantitative methods, was used. Also a number of data analysis procedures such as descriptive statistics and inferential statistics, were used the results of which will be fully described below.

To answer the first research question a corpus analysis was done. The number of spoken features such as time creation device, facilitation device and stock phrases was counted and the frequencies obtained were described in terms of percentage and frequencies. As table 2 shows, time creation devices and stock phrases are the two most frequently used strategies in the present study, and as it is shown in table 2, time creation devices are occupying the absolute majority of all strategies employed. Among time creation devices, the overuse of non-word meaningless fillers (such as uh, um, er, etc) which can be seen in both groups (high proficient group

used 52.60% of such meaningless fillers, with the frequency of 158 out of 300, and low proficient group used 51.66% of such fillers, with the frequency of 124 out of 239), one reason can be the context of interview of the oral test in which subjects tend to be unnaturally nervous or anxious; Our subjects have not been fully aware that apart from non-word fillers and repetitions, there are other means to win time without sacrificing fluency, for instance, speakers can ask for help from the other interlocutor, or use meaningful fillers like "How should I put it", and so on. Generally, subjects employ far more frequently strategies of non-word meaningless fillers, time creation fillers and repetition than they employ strategy of facilitation and fixed device or other achievement strategies.

Table 1

Total number of Words

	Proficiency Level	Total
Total Number of Words	High-proficient	6980
	Low-proficient	5800

Table 2

The frequency and percentage of spoken features

Spoken features	Time creation devices		Facilitation devices		Stock phrases	
	frequency	percentage	frequency	percentage	frequency	percentage
High-proficient	300	4.29%	15	0.21%	38	0.54%
Low-proficient	240	4.13%	5	0.08%	27	0.46%

Furthermore, subjects used some specific time creation fillers more than others. For example, high-proficient group used 17.33% for "I think", 11.33% for "you know" and "as you know" 5.33% for "I mean" and 5% for "well". Low-proficient group used 16.25% for "I think", 13.75% for "you know" and "as you know", 4.58% for "I mean" and 2.5% for "well". An interesting point is that low-proficient group used 3.33% for "I do not know" which high proficient group did not use at all. One reason can be the proficiency level, since they cannot express themselves in English well, they use such fillers. Also based on our observation subjects respond to our prompts during the time of interview revealed that they suffer from lack of self-confidence.

As results indicate, high-proficient group used fillers like "something like that" or "something like this", for generic words which are very common in spoken interaction. Even though they are not precise, they are fully comprehensible in the speaking situation because they talk about people, things or activities that can be seen or because they are familiar to the speakers. They make spoken communication quick and easy, and few people would find anything strange about this in their mother tongue. Generic words may also come naturally to second-language learners, but in a foreign language context where learners have few opportunities to speak the language outside the classroom; this feature of spoken language may be harder to notice and learn (Luoma, 2004). Therefore, it is clear that why only high-proficient group used such generic words in this study, in fact, it is rather hard for the EFL learners to learn and use such time creation fillers appropriately in their speech production.

As the results of the study indicate and as Luoma (2004) states, EFL learners use features of real-world listening input such as time creation device (pause fillers), facilitation device (fixed and conversational phrases), stock phrases and compensation device such as redundancy (repetition, reformulation, and rephrasing) in their speech production. The more students face with appropriate input, the better they use such spoken features. Therefore, teachers should provide students with enough input to enhance their output.

For the second part, to see whether there is any correlation between the participants' use of these features with their proficiency levels, an independent t-test sample was calculated. According to t-score, it can be concluded that statistically there is no significant difference between two groups regarding their proficiency levels and frequency of appropriate use of spoken features (table 3).

Table 3*T-score results*

	Mean	SD	t-score
High-proficient	20.00	40.76	0.32
Low-proficient	15.66	32.12	

Note. DF =28 $p < 0.05$

The result of the quantitative phase reveals the fact that EFL learners need to be taught more communicative strategies and skills to enhance their speaking ability. In language learning, we are moving from mastering the language form to using the language communicatively, therefore, learners should be presented with more authentic input and appropriate tasks in order to speak and communicate successfully.

Moreover, "that comprehension is ahead of production is a linguistic universal of acquisition..." (Ingram, 1974) and "speaking, actual production, does nothing to directly cause second language acquisition, since it is only input that counts. Speaking is the result of acquisition. The ability to speak a second language emerges or develops on its own after the acquirer has built up enough competence listening and reading" (Krashen, 1984). Also as it can be seen in Ferguson's (1998) work, without production being developed, comprehension will not develop. The fact remains then, we speak better than we understand. Therefore, to improve the speaking skill of learners, teachers first need to practice other skills to pave the ground for acquiring speaking.

6. Conclusion

The Speaking skill is viewed as the most substantial part of an EFL course. The growing need for international communication in the information age, has led many language learners to language classes in order to improve their speaking ability. Even though many students have mastered basic speaking skills, some students have performed much more effectively in the acquisition and their progress in oral communication than others. And those who are more effective communicators experience more success in school and in other areas of their lives (Malmir & Shoorcheh, 2012). According to Folse (2006), for most people, being able to speak a language means knowing that language since speech is the most fundamental means of communication among human beings. Nevertheless, speaking in a second or foreign language learning context has often been viewed as the most demanding of the four language skills. Speaking a language is especially difficult for foreign language learners because effective oral communication requires the ability to use the language appropriately in social interactions (Fulcher, 2003). Bygate (1987) declared that to become a proficient EFL speaker, studying the knowledge of grammar, vocabulary, pronunciation, intonation, etc. is not adequate but the ability to use this knowledge in order to communicate successfully is indispensable. From a testing perspective, as noted by Luoma:

speaking is special because of its interactive nature. It is often tested in live interaction, where the test discourse is not entirely predictable, just as no two conversations are ever exactly the same even if they are about the same topic and the speakers have the same roles and aims (Luoma, 2004, p. 170).

Speaking skills are important part of the curriculum in language teaching as well, and this makes them an important object of assessment as well (Luoma, 2004, p. 1). Assessing speaking is challenging, and as mentioned before, Speaking is a meaningful interaction between people. Some applied linguistic analyses, especially ones that focus on linguistic features, may at first sight seem to emphasize form at the cost of meaning, but in fact the concepts that are used in them are closely related to meaning as well. An important point from the linguistic description of spoken language is the special nature of spoken grammar and spoken vocabulary. This is especially important in creating rating criteria. Regarding interactive and social features of speaking, few people would question their relevance for understanding what speaking is like. From our experience with speaking, we know that conversations with different people turn out to be different even if we ourselves have more or less the same things to say, because speakers react to each other and construct discussions together (Luoma, 2004).

Then the aim of English speaking class is to enable students using the target language as social functions, to express their ideas, to handle basic interactive skills, as well as to present their needs, such as expressing opinions. For these purposes, English teachers might design a construction syllabus in which students first develop simple conversation skills, and then build on those skills in order to further achieve more complex skills (Brown & Yule, 1983). Moreover the teachers should employ a set of strategies to help the student's performance progress.

Therefore, EFL learners need to learn words, phrases and strategies to communicate. These expressions and strategies are very common in native speaker speech, but rather difficult for foreign language learners to acquire. As the results of this study show, even high-proficient learners use far more frequently non-word meaningless fillers to create time to speak which is overused than other strategies. Teachers should encourage students to learn and use other communication strategies as well.

6.1 Pedagogical Implications

The results obtained in this study would suggest a number of implications which can be useful for teachers, testers and language practitioners in EFL context. First and foremost, when language teachers are designing tasks for class activities, they should keep an eye to the variation in the level of difficulty so as to meet different needs of the students, and the task should also be interesting and realistic to students so that students could have free association for talking and strong motivation for interaction. Another problem in EFL context and more specifically in the Iranian EFL context, as noted by Mohammadzadeh, Dabaghi, and Tavakoli (2013), is that many teachers might avoid using a task-based approach on the account of the fact that it does not provide enough opportunity for language learners to focus on their speaking.

Non-word meaningless fillers and repetition are not the best achievement strategies in filling the gap between inter-language competence and intended communication goal in L2. Teachers should encourage students to learn and skillfully use other achievement strategies like reformulation. When reduction strategies are unavoidable, students can be taught to incorporate some easy-to-access achievement strategies in a view to mitigating the abruptness of shifting topic or the embarrassment of remaining salience for a long time. It should bear in mind that Fluency should only be encouraged at no sacrifice of accuracy.

Implications arising from this study suggest that teachers, syllabus designers and material developers, as mentioned by Assia and Said (2014), should take into account several important principles of language learning such as learner autonomy, learner needs and interests, learner involvement, cooperation and interaction among learners, self and peer-assessment and this is through including interactive activities that provide students with a sense of achievement and personal accountability and encourage them to think about the process of language learning and how to approach it more effectively.

6.2 Limitations and directions for further research

As for the limitations of the study, first, it can be said that the findings presented in this paper is the results of a classroom experience, therefore, the inferences drawn from this study is limited due to the nature of the participants. The participants of the study were thirty four junior EFL students at the University of Isfahan. Other researches with larger sample size and with students of various proficiency levels at different universities can be done to enhance the generalizability of the findings. Then, more data is needed, using different quantitative and qualitative research methods. Since oral communication fulfills a number of general and discipline-specific pedagogical functions, there are several areas of potential research. The interaction of age, gender and speaking assessment that was left untouched in the present study is one of the potential areas that can handled in further research.

7. References:

- Assia, B., & Said, K. (2014). The impact of small group interaction on learners' grammatical accuracy achievement. *International Journal of Research Studies in Language Learning*, 3(2), 85-100. <http://dx.doi.org/10.5861/ijrsl.2013.511>
- Bardovi-Harlig, K. (1999). Exploring the interlanguage of interlanguage pragmatics: A research agenda for acquisitional pragmatics. *Language Learning*, 49, 677-713. <http://dx.doi.org/10.1111/0023-8333.00105>
- Bonk, W. J., & Ockey, G. J. (2003). A many-facet Rasch analysis of the second language group oral discussion task. *Language Testing*, 20(1), 89-110. <http://dx.doi.org/10.1191/0265532203lt245oa>
- Brown, A. (2000). *An Investigation of the rating process in the IELTS speaking module*. In R. Tulloh (Ed.), IELTS research reports 1996, (Vol. 3, pp. 49-85) Sydney: ELICOS.
- Brown, A. (2003). Interviewer variation and the co-construction of speaking Proficiency. *Language Testing*, 20(1) 1-25. <http://dx.doi.org/10.1191/0265532203lt242oa>
- Brown, G. (1996). Language learning, competence and performance. In G. Brown, K. Malmkjaer, & J. Williams (Eds.), *Performance and competence in second language acquisition* (pp. 187-203). Cambridge: CUP.
- Brown, G., & Yule, G. (1983). *Teaching the spoken language: An approach based on the analysis of conversational English*. Cambridge: CUP.
- Burns, A., & Seidlhofer, B. (2002). Speaking and pronunciation. In N. Schmitt (Ed.), *An introduction to applied linguistics*. London: Arnold.
- Bygate, M. (1987). *Speaking*. Oxford: Oxford University Press.
- Celce-Murcia, M., & Olshtain, E. (2000). *Discourse and context in language teaching*. New York: Cambridge University Press.
- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: CUP.
- Ellis, N. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24, 143-188. <http://dx.doi.org/10.1017/S0272263102002024>
- Farhady, H., Jafarpour, A., & Birjandi, P. (1994). *Testing language skills*. SAMT Publications.
- Ferguson, N. (1998). Comprehension and production of the spoken language. *International Review of Applied Linguistics in Language Teaching*, 34(4), 307-322.
- Folse, K. (2006). *The art of teaching speaking*. Michigan: Michigan University Press.
- Freed, B. (1995). What makes us think that students who study abroad become fluent? In B. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 123-148). Amsterdam: John Benjamins. <http://dx.doi.org/10.1075/sibil.9.09fre>
- Fulcher, G. (1996). Does thick description lead to smart tests? A data-based approach to rating scale construction. *Language Testing*, 13, 208 - 238. <http://dx.doi.org/10.1177/026553229601300205>
- Fulcher, G. (2003). *Testing second language speaking*. New York: Pearson Longman.
- Harris, D. P. (1969). *Testing English as a second language*. New York: MacGraw Hill Book Company.
- Hartley, L., & Sporing, M. (1999). Teaching communicatively: Assessing communicatively? *Language Learning Journal*, 19, 73-79. <http://dx.doi.org/10.1080/09571739985200141>
- Hasselgren, A. (1998). *Smallwords and valid testing*. Unpublished doctoral thesis. Department of English, University of Bergen, Bergen, Norway.
- House, J. (1996). Developing pragmatic fluency in English as a foreign language: Routines and meta-pragmatics awareness. *Studies in Second Language Acquisition*, 18, 225-252. <http://dx.doi.org/10.1017/S0272263100014893>
- Hughes, A. (2003). *Testing for language teachers*. Cambridge: Cambridge University Press.
- Ingram, D. (1974). *The Relationship between comprehension and production in language perspectives*. Baltimore University Press.
- Kasper, G. (1996). Introduction: Interlanguage pragmatics in SLA. *Studies in Second Language Acquisition*, 18, 145-148. <http://dx.doi.org/10.1017/S0272263100014856>

- Koponen, M. (1995). Let your language and thoughts flow! Is there a case for fluency in ELT and applied linguistics? Paper presented at the *Language Testing Forum, Newcastle, UK, November 1995*.
- Krashen, S. (1984). Immersion, why it works and what it has taught us. *Language and Society, 12*(1).
- Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning, 3*, 387–417. <http://dx.doi.org/10.1111/j.1467-1770.1990.tb00669.x>
- Luoma, S. (2004). *Assessing speaking*. Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511733017>
- Malmir, A., & Shoorcheh, S. (2012). An investigation of the impact of teaching critical thinking on the Iranian EFL learners' speaking skill. *Journal of Language Teaching and Research, 3*(4), 608-617. <http://dx.doi.org/10.4304/jltr.3.4.608-617>
- McNamara, T. (1996). *Measuring second language performance*. London: Longman.
- Mohammadzadeh, A. M., Dabaghi, A., & Tavakoli, M. (2013). The effects of simultaneous use of pre-planning along +/-Here-and-Now dimension on fluency, complexity, and accuracy of Iranian EFL learners' written performance. *International Journal of Research Studies in Language Learning, 2*(3), 49-65. <http://dx.doi.org/10.5861/ijrsl.2012.168>
- Nattinger, J., & DeCarrico, J. (1992). *Lexical phrases and language teaching*. Oxford: OUP.
- Oller, J. W., & Perkins, K. (1977). *Research in language testing*. Newbury House Publishers, Inc.
- Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Native like selection and native like fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and Communication*. London: Longman.

