

TEACHING AND ASSESSING SPEAKING

Ahmet ÖNAL

1. Introduction

The integral role played by the skill of speaking in the process of communication cannot be overstated and it can be argued that we speak more than we write. However, most of the learners especially in English as a foreign language contexts (including Turkey) state that *they can understand but they cannot speak* when it comes to communicating in English. Accordingly, this chapter aims to provide pre-service and in-service English language teachers with some useful ideas for teaching and testing the skill of speaking. More specifically, the chapter outlines the process and different types of speaking, identifies the difficulties encountered when teaching/testing speaking and offers solutions, reviews alternative ways of testing speaking and focuses on the use of rubrics and the responsibility of the test designers and raters throughout the process of assessment of speaking.

Traditional view of second/ foreign language teaching literature divides four language skills into two; namely, listening and reading are regarded as *receptive* skills whereas speaking and writing are considered as *productive* skills since users of a language are expected to produce written or spoken texts as a consequence of their performance. Although both writing and speaking are seen as productive skills, differences between speaking and writing have been highlighted by many researchers (Louma, 2004; Nunan, 2015; Richards, 2015). For instance, while writing is primarily based on sentences, speech is more organized around idea units (Louma, 2004). We make use of punctuation to organize our writing; however, in speech, intonation, pitch and stress substitute punctuation. In writing, we have the chance and time to modify and improve our product until we feel satisfied with it in most cases; however, speaking is similar to first draft writing and we rarely have the chance to do a second draft (Nunan, 2015). This inventory of differences can be extended even to the people addressed. When we write, most of the time, we do not see the instant reactions of our readers; however, when we speak to some people (as long as it is an interaction that involves at least two parties), we generally get some clue from the facial expressions of the interlocutors even if they do not respond to our message orally.

2. What does Speaking Involve?

It has been observed by Richards (2015) that learners take hundreds of hours of English courses, learn (or, in some cases, *memorize*) hundreds or even thousands of new words and/or grammatical structures and sit for and pass many written – often multiple choice – exams; however, they tend to evaluate their progress and the efficiency of the English program on the basis of how much they have improved their speaking skills. Thus, it has been argued that knowing a language is assumed as being able to speak that language (Nunan, 2015) and the crucial role played by speaking in the whole process of communication has been frequently highlighted (Richards, 2015; Yaman & Özcan, 2015). A judicious command of vocabulary and grammar is essential, but not sufficient for effective communication (Nunan, 2015) and such other competences as sociolinguistic, discourse and strategic competences, phonological skills,

speech function, interactional skills, extended discourse skills, and conversation management strategies are also needed (Bohlke, 2014; Goh, 2007; Szymańska-Czaplak, 2015; Yaman & Özcan, 2015).

More precisely, the act of speaking involves, first of all, having an opinion to express; and then, finding the appropriate words and organizing them into units in one's mind. Speech organs should be employed to articulate and convey the message and it should be constantly monitored and corrected if any mistakes are committed. The speaker should also be aware of social, cultural, sociolinguistic, pragmatic and discursive conventions while planning and delivering his/her message (Watkins, 2005), which further complicates the whole process. Put differently, as has been put forward by Levelt (1993), four stages (conceptualization, formulation, articulation and self-monitoring) are essential for oral production. The speaker, first of all, needs to plan or *conceptualize* the content of his/her message. The content should be *formulated* in appropriate words and/or phrases in line with grammatical conventions. Then, the message should be physically produced or *articulated* using the speech organs and the speaker needs to control the whole process with the help of his/her *self-monitor*. This four-stage process may seem extremely laborious and time-consuming; nevertheless, as the speaker practices and gains automaticity, his/her fluency as well as accuracy will also develop (Harmer, 2007).

3. Types of Speaking

The skill of speaking has traditionally been regarded as an individual entity in itself; however, the Common European Framework of Reference for Languages (CEFR) draws a line between *productive* and *interactive* speaking (Council of Europe, 2001). Interactive speaking refers to spontaneous exchange characteristic of conversation whereas productive speaking involves preplanned and/or rehearsed, mostly monologue type of speech (Green, 2021). In this respect, in interactive speaking tasks, the speakers should also be effective listeners and process the input so that they can construct their own output (Galaczi & Taylor, 2021). In a similar vein, Nunan (2015) distinguishes between *reproductive* speaking, where the speaker repeats or imitates the model form s/he hears, and *creative* speaking, where speakers construct and produce their original ideas. Similarly, Louma (2004) divides speaking tasks as *structured* (in which what the learners say is precisely specified) and *open-ended* (in which various ways of fulfilling the task is allowed). According to Nunan (2015), teachers need to make the best use of both types of speaking by arranging the proportion with a specific view to the context (including the interests, aims, needs, proficiency levels of the learners) in which they function.

It should not go without saying that the skill of speaking comprises of a wide variety of *genres*, defined as "...knowledge of different types of spoken interaction, including the discourse conventions of each kind of interaction, as well as the sociocultural and pragmatic dimensions of different genres" (Richards, 2015, p. 408). Effective speakers need to engage in a diversity of genres from small talk and casual conversations to debates and transactions when communicating in their daily lives; therefore, a wide range of speaking activities that learners

may engage in their real lives need to be offered to the learners (Nunan, 2015; Watkins, 2005; Wigglesworth & Frost, 2017). To exemplify, learners should be required to describe, narrate, instruct, inform, explain, complain, persuade, predict, decide, compare and/or contrast when speaking because people speak with such aims in mind in real life (Louma, 2004). It should also be noted that some individuals are naturally good at, let's say, narrating, but not so good at complaining. Therefore, providing them with a variety of speech acts, types and genres will possibly help them improve their weaknesses.

4. Activities to Teach / Practice Speaking

The ideas presented in this section of the chapter should be taken as suggestions based on the review of the relevant literature since the array of activities that can be utilized when teaching or practicing speaking is only limited to the creativity and imagination of the teacher. To start with, it has been recommended that activities that feature an *information gap* should be employed because they promise to generate real conversation by encouraging negotiation of meaning, which is regarded as vital for language acquisition (Harmer, 2007; Nunan, 2015). It should also be noted that speaking lessons that encourage creativity provide the learners with an ideal and protected environment, which enables them to improve their self-confidence (Watkins, 2005) and learners at lower levels (CEFR A1 – B1) may need verbal or visual interactional support (or *scaffolding*) when they are asked to speak (Galaczi & Taylor, 2021). As to the dichotomy between *fluency* and *accuracy*, it has been asserted that fluent but inaccurate, or accurate but disfluent speakers cannot be regarded as effective communicators; hence, teachers should try to provide the learners with practice for both fluency and accuracy (Nunan, 2015). Furthermore, *integration* of four skills in the processes of both teaching and testing has been strongly advocated since, in our real lives, we rarely use one skill in isolation (Brown & Abeywickrama, 2018; Harmer, 2007; Louma, 2004). Likewise, input on which the learners may construct their output should be presented in various forms and modalities and teachers need to try to make the best use of the latest innovations in technology (Nakatsuhara et al., 2022). As an example, the topic for a discussion activity may be related to a video that learners have been assigned to watch prior to the class. Likewise, web 2.0 tools may be effectively utilized to provide the learners with the necessary input and encourage oral production.

Traditional activities such as reading aloud, sentence repetition/completion, ordering activities and cued dialogues – though lacking creativity and information-gap – may be employed to practice speaking. Similarly, picture description, question and answer, story-(re)telling, personal recounts, class questionnaires/surveys, 'find the difference' pictures may also be used as speaking practice activities (Baker & Westrup, 2000; Harmer, 2007; Louma, 2004; Nakatsuhara et al., 2021; Nakatsuhara et al., 2022; Richards, 2015). Elicited imitation tasks, in which learners are asked to read aloud written sentences or listen to utterances and repeat them, have recently regained their popularity as well, as a response to the developments in automated speech recognition and evaluation technology (Nakatsuhara et al., 2022). Furthermore, problem

solving activities, oral presentations, role plays, simulations, reacting in given situations, (group/pair) discussions, games, competitions, interviews, moral dilemmas and debates may be used to encourage and improve learners' speaking skills (Brown & Abeywickrama, 2018; Galaczi & Taylor, 2021; Harmer, 2007; Louma, 2004; Watkins, 2005).

As has been reported by Nunan (2015), more than half of the class time is occupied by teacher talk; however, speaking is learnt and improved by speaking, which implies that students' talking time needs to be maximized. One way of achieving this, especially in crowded classrooms, is to carefully organize pair- and group-work, which may contribute to the interactional competence of the learners by enhancing their skills in agenda, turn and topic management, non-verbal behavior (including eye contact, facial expression, posture, etc.), negotiation of meaning, turn-taking, back-channelling, speaker selection and change, among many others (Brown & Abeywickrama, 2018; Galaczi & Taylor, 2021; Nakatsuhara et al., 2022; Richards, 2015). Furthermore, as opposed to teacher-to-learner or learner-to-teacher interaction, such learner-to-learner interaction will be more real life-like as learners can be considered in equal power positions (Jankowska & Zielińska, 2015; Louma, 2004). It should be cautioned at this point that there are several drawbacks of pair-/group-work activities such as increased native language use, deviation from the actual activity or domination of the activity by an individual student; thus, teachers should closely monitor their students to ensure that each student is on task and offer scaffolding when needed. In this respect, the effect of the interlocutor has been regarded as a decisive factor for the success of the communication process because interlocutors need to collaborate and negotiate meaning throughout the process (Brown, 2012; Fulcher & Davidson, 2007; Gablasova, 2021; Galaczi & Taylor, 2021; Green, 2021; Louma, 2004; Nakatsuhara et al., 2022; Nunan, 1991; Ross, 2012). More precisely, specific features of a partner such as his/her gender, personality, appearance, accent, communication style and even language level may exert either constructive or destructive influence on the performance of his/her partner (Jankowska & Zielińska, 2015; Louma, 2004). Therefore, when pair-/group-work activities are to be employed for instructional and/or assessment purposes, the effects of the interlocutors need to be taken into account.

5. Problems in Teaching Speaking

The main problem in the teaching of speaking, according to Watkins (2005), is that when a lesson is focused on speaking, some learners may feel dissatisfied and some teachers may feel guilty due to not learning or teaching something with a clear and tangible outcome. In other words, some students cannot quite see the point of the class unless they take notes, deal with structures or do some rote memorization. In this respect, the task of the teacher is, first of all, to transform the classroom into a '*talking classroom*' by building a classroom culture of speaking (Harmer, 2007) because teaching a second/foreign language and language skill is quite different from teaching geography or physics. It is true that students can learn geography or physics by listening to lectures, taking notes, memorizing formulae and solving problems; in contrast, students can learn speaking in a second/foreign language only by speaking.

Additionally, the context in which the instruction is offered should be taken into consideration by the teachers because learners in a second language context have greater opportunities to get exposed to and use the target language when compared with their counterparts in a foreign language context (Nunan, 2015). Put differently, classroom is possibly the only place where they can get exposed to, practice and speak English for learners in an English as a Foreign Language (EFL) context, which has significant implications for the teachers.

Another point to consider is that speaking is generally regarded as the most anxiety-provoking aspect of second/foreign language education (Büyükkaracı, 2017; Cheng et al., 1999; Hewitt & Stephenson, 2012; Kitano, 2001; Marzec-Stawiarska, 2015) and some students may feel shy and anxious to speak (Watkins, 2005). A supportive and encouraging classroom atmosphere is essential for such students to overcome their negative feelings and mistakes committed by learners should be handled with great care so as not to discourage them. It has been observed by Louma (2004) that teachers tend to offer informal feedback for speaking, which is hardly organized and informative to the learners. However, useful feedback should be concrete, descriptive, instructive and constructive. As the focus is on fluency when teaching speaking, correction for the mistakes should be offered positively and without discouraging the learners (Baker & Westrup, 2000; Harmer, 2007; Nunan, 2015).

6. Assessment of Speaking

Before we start our discussion on *how* to assess speaking, a more central question awaits its answer: *Should we assess speaking?* The reason why this section of the chapter begins with such a question lies in the experiences and observations of the author of the chapter. Strictly speaking, some English language teachers in Turkish context (and possibly in many other contexts) simply do not assess their learners' speaking skills. Some of these teachers blame the strict and overloaded curriculum and/or textbooks mandated by the Ministry of National Education, crowded classrooms, tight schedule, lack of time and/or expertise, centrally organized (high-stakes) exams and even their students (as they lack the interest and motivation) for not testing speaking. They may be justified to some extent; yet, it would not be possible to improve learners' speaking skills and encourage positive washback unless the skill of speaking is tested (Fulcher & Davidson, 2007; Louma, 2004). This implies that English language teachers should certainly assess speaking whatever obstacles they encounter if they are to improve their learners' speaking skills.

The assessment of productive skills, in particular the skill of speaking, has been attached greater significance (Richards, 2015), yet it presents several challenges for the test designers and teachers since it is an expensive, time consuming, logistically-complex and resource-intensive endeavor (Brown & Abeywickrama, 2018; Hingle & Lington, 2002; Jankowska & Zielińska, 2015; Nakatsuhara et al., 2021; Önalın, 2020) and the best way of testing speaking is getting learners to speak directly via well-organized tasks (Fulcher & Davidson, 2007; Green, 2021; He & Young, 1998) since it features higher content validity (Brown & Abeywickrama, 2018).

Put differently, just asking the test-takers to *speak* is not the proper way of testing speaking because assessment of speaking also involves the assessment of pronunciation, intelligibility, fluency, grammar, vocabulary, coherence, cohesion and the ability to interact and adjust one's speech to a particular social context (Jankowska & Zielińska, 2015). Therefore, it would be justified to argue that the design, delivery and scoring of speaking tests demand a great deal of planning, effort, expertise, investment and time. As an example, it has been suggested that developers of speaking assessments should *a*) define the kind of speaking they want to test in a particular context; *b*) develop tasks and rating criteria that test this (including instructions and materials, pictures, role cards, etc.); *c*) inform the examinees about what they test; and *d*) make sure that the testing and rating processes actually follow the stated plans (Louma, 2004, pp. 28-29). This clearly shows that a paradigm shift is called for in the assessment practices and non-traditional, or more precisely, *alternative* ways of assessment need to be designed and adopted.

6.1. Modern ideas for assessment of speaking

Activities suggested for the teaching of speaking skill can also be effectively adapted and employed for the assessment of speaking. Considering the level of the learners and the purpose of the assessment, traditional techniques such as reading aloud, sentence repetition, cued dialogues, picture description and story-(re)telling, among a plethora of others, can be utilized (Harmer, 2007; Louma, 2004; Richards, 2015). On the other hand, techniques that embody information-gap and call for more creativity such as problem-solving, role plays, simulations, debates, discussions and interviews may also be used to test learners' speaking skills (Brown & Abeywickrama, 2018; Galaczi & Taylor, 2021; Harmer, 2007; Louma, 2004; Watkins, 2005). More specifically, oral interviews have been regarded as good examples of performance-based language assessment techniques (Brown & Abeywickrama, 2018) and they can be implemented face-to-face or, thanks to the innovations in technology, through web-conferencing programs or applications, which enable two or more users in different locations to see and hear each other and interact in real time via high-speed internet connection and mobile tools such as smartphones, tablets or laptops (Lim, 2018; Loranc-Paszyk, 2015; Louma, 2004; Nakatsuhara et al., 2021; Nakatsuhara et al., 2022). One big advantage of such technology is the possibility of recording the meetings, which allows the rater to watch and re-watch the performances of the learners and score them more accurately rather than relying on his/her memory as to the performance of a specific learner (Jankowska & Zielińska, 2015). During the global COVID-19 pandemic, almost all kinds of social, professional and educational interactions had to be conducted online and it seems highly likely that the conveniences offered by the internet and computer technology will not be discarded in the *new normal* period; thus, the use of web-conferencing tools for both teaching and testing speaking contributes to authenticity since it reflects real-life experiences of the learners (Nakatsuhara et al., 2022).

6.2. Alternative assessment of speaking

Traditional assessment usually deals with what learners know (declarative knowledge); however, alternative assessment focuses more on what learners can do (procedural knowledge) in the target language (Huerta-Macías, 2002; Ross, 2012). As has been reported by Brown and Abeywickrama (2018), a more student-centered agenda is kept by language teachers and testers across the globe and the use of alternative assessment methods such as portfolios, interviews, journals, project work, self- or peer-assessment, checklists of student behaviors or products, reading logs, videos of role plays and discussions, self-evaluation questionnaires, work samples, and teacher observations or anecdotal records has been suggested (Huerta-Macías, 2002; Jankowska & Zielińska, 2015; Louma, 2004; Richards & Renandya, 2002) because they give the learners the chance to control and regulate their own learning by involving them more in the process and placing them in the center (Jankowska & Zielińska, 2015; Richards, 2015; Richards & Renandya, 2002). In other words, students become more *autonomous* and develop more engagement in their own destiny (Brown & Abeywickrama, 2018). Peer-assessment, for instance, encourages learners to learn from and collaborate with each other by paying more attention to what their peers are saying when they are not speaking themselves (Jankowska & Zielińska, 2015; Louma, 2004). As a result, teachers get the chance to save some time and effort by sharing the responsibility of giving feedback. It should be highlighted that this does not come to mean that self- or peer-assessment should replace teacher assessment but they can certainly supplement it.

6.3. Rubric use in the assessment of speaking

For the assessment of performance, a *scoring rubric* (or a rating scale) that includes the criteria upon which the scoring is based is essential (Brown & Abeywickrama, 2018; Fulcher, 2012; Galaczi & Lim, 2022; Kuiken & Vedder, 2021; Richards, 2015; Wigglesworth & Frost, 2017; Xi, 2012) although many other less reliable ways of scoring performance such as *impression scoring* (in which an arbitrary score based on the highly subjective judgment of the rater is assigned to performances), *error counts* (in which points are deducted for each error in the performance) and *checklists* (which include descriptors of the desired level of performance that need to be ticked or crossed by the rater) have also been utilized (Green, 2021). The main problem with impression scoring is its highly subjective (and thus, unreliable) nature. Similarly, learners tend to avoid taking risks by producing shorter and simpler sentences so as not to make mistakes if they know that their errors are counted. Checklists, on the other hand, can be regarded as precursors of rubrics as they also involve performance descriptors; however, checklists do not allow the raters to grade or weigh the descriptors in line with the requirements of the test purpose and context. At this point, the design and wording of the descriptors has been addressed by many researchers in that they should be brief, clear, explicit, concrete, practical, consistent, transparent and standards-based (Brown & Abeywickrama, 2018; Council of Europe, 2001; Kuiken & Vedder, 2021).

Owing to the problems associated with impression scoring, error counts and checklists, raters and test-designers more commonly employ holistic and analytical scoring rubrics. In *holistic* scoring, a single number is used to score a speech sample and that specific number is a generalization of the performance of the test-taker on many other possible speaking tasks (Fulcher, 2012; Fulcher & Davidson, 2007; Green, 2021; Kuiken & Vedder, 2021; Louma, 2004; Richards, 2015). On the other hand, raters are required to award a certain number of different scores distributed and weighted over a range of criteria in *analytical* scoring (Brown & Abeywickrama, 2018; Galaczi & Lim, 2022; Green, 2021; Richards, 2015). It has been reported that holistic scoring is easy to use, practical and produces higher inter-rater reliability at the cost of validity (Galaczi & Lim, 2022; Louma, 2004; Weigle, 2002). It has also been noted that holistic scoring cannot provide the learners with correction, feedback, or diagnosis as to their individual performances, which is possible with analytical scoring (Brown & Abeywickrama, 2018; Hamp-Lyons, 1991; Kuiken & Vedder, 2021; Louma, 2004). Although it takes longer to score the performance of test-takers using an analytical rubric, it is believed that raters can understand and apply the criteria more accurately since the descriptors are designed in a more detailed fashion (Kuiken & Vedder, 2021). As can be seen, both holistic and analytical scoring have their pros and cons. Accordingly, Fulcher (2012) argues that the purpose and context of assessment need to be taken into account when deciding on the scoring approach to be employed since “...one scoring approach is not inherently superior to another” (Galaczi & Lim, 2022, p. 500).

Apart from the type of the scoring rubric employed, *reliability* emerges as an important aspect to be considered in the assessment of speaking and qualified raters as well as high-quality scoring instruments and procedures are essential for ensuring high reliability (Louma, 2004). When two (or more) raters are scoring the same performance, the degree to which the raters agree with each other is referred to as *inter-rater reliability* and the degree to which an individual rater agrees with him/herself when scoring the same performance on different occasions is defined as *intra-rater reliability* (Fulcher & Davidson, 2007; Louma, 2004). In tests that require subjective judgment of the raters, high level of intra-rater and inter-rater reliability is desirable since “...it should be *a matter of indifference to a test taker who scores the performance*” (Fulcher & Davidson, 2007, p. 131) and one rater should be interchangeable with another rater from the perspective of the test-taker (Brown, 2012).

6.4. Rater training for the assessment of speaking

In many types of speaking tests, the rater needs to fulfill a dual task; s/he has to act as an interlocutor and score the performance of the test-taker, which is a challenging undertaking and which may result in the rather undesirable *rater effect*. The rater effect, which involves variables such as raters’ decision-making styles, demographic backgrounds, training and experience as well as test-takers’ acquaintanceship, gender, extroversion, talkativeness, cultural background, and language proficiency should also be paid precise attention in performance assessment and needs to be minimized to achieve higher reliability (Brown, 2012; Brown & Abeywickrama,

2018; Davis, 2022; Galaczi & Lim, 2022; Galaczi & Taylor, 2021; Nakatsuhara et al., 2022; Ross, 2012; Wigglesworth & Frost, 2017). As can be inferred, the assessment of performance calls for qualified and trained personnel since just asking the test-takers to speak or write will not work and the performance must be elicited through well-designed tests and capable raters (Brown, 2012; Louma, 2004). Accordingly, the significance and necessity of regular training of all the staff, in particular the raters, responsible for the administration of the test and achieving standardization has been repeatedly highlighted (Brown, 2012; Brown & Abeywickrama, 2018; Davis, 2022; Fulcher & Davidson, 2007; Gablasova, 2021; Galaczi & Lim, 2022; Green, 2021; Hamp-Lyons, 1991; Kuiken & Vedder, 2021; Louma, 2004; Lumley, 2002; Ross, 2012; Wigglesworth & Frost, 2017) since it is the raters who mediate between the rubric and the performance by paying more attention to certain dimensions of the rubric and the specific performance (Pill & Smart, 2021). It has also been noted that such rater training programs can be conducted online so that logistical barriers can be overcome (Brown, 2012).

7. Conclusion

Although there are a multitude ways of communication among humans (such as writing, sign languages, use of body language, pictures and signs, etc.), speaking has always functioned as the basic way of interaction. This implies that due importance needs to be given to both teaching and testing the skill of speaking. To do this, first of all, the nature of the speaking skill should be clearly defined and understood by the teachers. They should be able design and deliver speaking tasks that feature authenticity and information-gap, encourage learners' creativity and imagination and involve real communication while both teaching and testing speaking. In addition, these tasks need to include variety since, as has been asserted by Green (2021) and Nakatsuhara et al. (2022), there is no perfect assessment task or tool that can be employed in any context and for any purpose; therefore, it would be wise to provide our learners with a wide variety of assessment tasks and tools in order to offset the weaknesses of a specific task or tool via the strengths of another task or tool. It is true that teaching and testing of speaking is a demanding undertaking that requires plenty of time, effort, investment and commitment; however, there is no other way of achieving this.

It should also be noted that we are living in the age of technology and today's students (labeled as *digital natives*) are continuously immersed in the latest technological tools. Teachers should also try to make the best use of the technology to overcome some of the problems that accompany the attempts to teach and assess speaking. For instance, Nakatsuhara et al. (2021) draw attention to the innovations in the fields of speech science, automatic speech recognition, and deep neural networking technologies, which have the potential to assist teachers in the teaching and assessment of speaking and suggest that, though the innovations are not yet mature enough to substitute us, they can support, facilitate and supplement us in designing, implementing and scoring tests (Brown, 2012; Galaczi & Lim, 2022; Lim, 2018). Therefore, it may be possible to teach and test speaking without exerting too much effort and spending excessive time with the help of such innovations.

REFERENCES

- Baker, J. & Westrup, H. (2000). *English language teacher's handbook: How to teach large classes with few resources*. Bloomsbury Publishing.
- Bohlke, D. (2014). Fluency-oriented second language teaching. In Celce-Murcia, M., Brinton, D. & Snow, M. A. (Eds.), *Teaching English as a Second or Foreign Language* (4th Edition), 121-135. National Geographic Learning.
- Brown, A. (2012). Interlocutor and rater training, in Fulcher, G. & Davidson, F. (Eds.), *The Routledge handbook of language testing*, 413-425, Routledge.
- Brown, H. D. & Abeywickrama, P. (2018). *Language assessment: Principles and classroom practices* (3rd Ed.). Pearson Education.
- Büyükkarcı, O. (2017). Türk öğrencilerin yabancı dil öğreniminde karşılaştığı güçlükler, in Balcı, Ö., Çolak, F. & Ekizer, F. N. (Eds.), *Her yönüyle yabancı dil üzerine araştırmalar*, 123-136. Çizgi Kitabevi Yayınları.
- Cheng, Y. S., Horwitz, E. K. & Schallert, D. L. (1999). Language anxiety: Differentiating writing and speaking components. *Language Learning*, 49(3), 417-446.
- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge University Press.
- Davis, L. (2022). Rater and interlocutor training, in Fulcher, G. & Harding, L. (Eds.), *The Routledge handbook of language testing* (2nd Ed.), 322-338, Routledge.
- Fulcher, G. (2012). Scoring performance tests, in Fulcher, G. & Davidson, F. (Eds.), *The Routledge handbook of language testing*, 378-392, Routledge.
- Fulcher, G. & Davidson, F. (2007). *Language testing and assessment: An advanced resource book*. Routledge.
- Gablasova, D. (2021). Corpora for second language assessments, in Winke, P. & Brunfaut, T. (Eds.), *The Routledge handbook of second language acquisition and language testing*, 45-53, Routledge.
- Galaczi, E. D. & Lim, G. S. (2022). Scoring performance tests, in Fulcher, G. & Harding, L. (Eds.), *The Routledge handbook of language testing* (2nd Ed.), 495-510, Routledge.
- Galaczi, E. D. & Taylor, L. B. (2021). Measuring interactional competence, in Winke, P. & Brunfaut, T. (Eds.), *The Routledge handbook of second language acquisition and language testing*, 338-348, Routledge.
- Goh, C. C. M. (2007). Teaching speaking in the language classroom. In Renandya, W. & Richards, J.C. (Eds.), *RELC Portfolio Series*, 1-12. SEAMEO Regional English Language Centre.

- Green, A. (2021). *Exploring language assessment and testing (2nd Ed.)*. Routledge.
- Hamp-Lyons, L. (1991). Scoring procedures for ESL contexts. In Hamp-Lyons, L. (Ed.), *Assessing second language writing in academic contexts*, 241–276. Ablex Publishing Corporation.
- Harmer, J. (2007). *How to teach English*. Pearson Education Limited.
- He, A. W. & Young, R. (1998). Language proficiency interviews: A discourse approach. In Young, R. & He, A. W. (Eds.), *Talking and testing: Discourse approaches to the assessment of oral proficiency*, 1–24. Benjamins.
- Hewitt, E. & Stephenson, J. (2012). Foreign language anxiety and oral exam performance: A replication of Phillips's MLJ study. *The Modern Language Journal*, 96(2), 170-189.
- Hingle I. & Linington, V. (2002). English proficiency test: The oral component of a primary school, in Richards, J. C. & Renandya, W. A. (Eds.), *Methodology in language teaching: An anthology of current practice*, 354-360, Cambridge University Press.
- Huerta-Macías, A. (2002). Alternative assessment: Responses to commonly asked questions, in Richards, J. C. & Renandya, W. A. (Eds.), *Methodology in language teaching: An anthology of current practice*, 338-343. Cambridge University Press.
- Jankowska, A. & Zielińska, U. (2015). Designing a self-assessment instrument for developing the speaking skill at the advanced level, in Pawlak, M. & Waniek-Klimczak, E. (Eds.), *Issues in teaching, learning and testing speaking in a second language*, 251-265, Springer.
- Kitano, K. (2001). Anxiety in the college Japanese language classroom. *The Modern Language Journal*, 85(4), 549-566.
- Kuiken, F. & Vedder, I. (2021). Scoring approaches: Scales/rubrics, in Winke, P. & Brunfaut, T. (Eds.), *The Routledge handbook of second language acquisition and language testing*, 125-134, Routledge.
- Levelt, W. J. (1993). *Speaking: From intention to articulation*. MIT Press.
- Lim, G. S. (2018). Conceptualizing and operationalizing second language speaking assessment: Updating the construct for a new century, *Language Assessment Quarterly*, 15(3), 215-218, <https://doi.org/10.1080/15434303.2018.1482493>
- Loranc-Paszyk, B. (2015). Videoconferencing as a tool for developing speaking skills, in Pawlak, M. & Waniek-Klimczak, E. (Eds.), *Issues in teaching, learning and testing speaking in a second language*, 189-203, Springer.
- Lumley, T. (2002). Assessment criteria in a large-scale writing test: What do they really mean to the raters? *Language Testing*, 19(3), 246–276. <https://doi.org/10.1191/0265532202lt230oa>

- Louma, S. (2004). *Assessing speaking*. Cambridge University Press.
- Marzec-Stawiarska, M. (2015). Investigating foreign language speaking anxiety among advanced learners of English, in Pawlak, M. & Waniek-Klimczak, E. (Eds.), *Issues in teaching, learning and testing speaking in a second language*, 103-119, Springer.
- Nakatsuhara, F., Inoue, C. & Khabbazbashi, N. (2021). Measuring L2 speaking, in Winke, P. & Brunfaut, T. (Eds.), *The Routledge handbook of second language acquisition and language testing*, 285-294, Routledge.
- Nakatsuhara, F., Khabbazbashi, N. & Inoue, C. (2022). Assessing speaking, in Fulcher, G. & Harding, L. (Eds.), *The Routledge handbook of language testing (2nd Ed.)*, 209-222, Routledge.
- Nunan, D. (1991). *Language teaching methodology: A textbook for teachers*. Prentice Hall.
- Nunan, D. (2015). *Teaching English to speakers of other languages: An introduction*. Routledge.
- Önalın, O. (2020). Automated scoring of productive skills in language assessment, in Çelik, S. (Ed.), *Preparing teachers for a changing world: Contemporary issues in EFL education*, 103-114, Vizetek.
- Pill, J. & Smart, C. (2021). Raters: Behavior and training, in Winke, P. & Brunfaut, T. (Eds.), *The Routledge handbook of second language acquisition and language testing*, 135-144, Routledge.
- Richards, J. C. (2015). *Key issues in language teaching*. Cambridge University Press.
- Richards, J. C. & Renandya, W. A. (2002). Assessment, in Richards, J. C. & Renandya, W. A. (Eds.), *Methodology in language teaching: An anthology of current practice*, 335-337. Cambridge University Press.
- Ross, S. J. (2012). Claims, evidence, and inference in performance assessment, in Fulcher, G. & Davidson, F. (Eds.), *The Routledge handbook of language testing*, 223-233, Routledge.
- Szymańska-Czaplak, E. (2015). Developing and testing speaking skills in academic discourse, in Pawlak, M. & Waniek-Klimczak, E. (Eds.), *Issues in teaching, learning and testing speaking in a second language*, 233-250, Springer.
- Watkins, P. (2005). *Learning to teach English: A practical introduction for new teachers*. DELTA publishing.
- Weigle, S. C. (2002). *Assessing writing*. Cambridge University Press.
- Wigglesworth, G. & Frost, K. (2017). Task and performance-based assessment, in Shohamy, E., Or, I. G. & May, S. (Eds.), *Language testing and assessment (3rd Ed.)*, 121-134, Springer International Publishing.

- Xi, X. (2012). Validity and the automated scoring of performance tests, in Fulcher, G. & Davidson, F. (Eds.), *The Routledge handbook of language testing*, 438-451, Routledge.
- Yaman, Ş. & Özcan, M. (2015). Oral communication strategies used by Turkish students learning English as a foreign language, in Pawlak, M. & Waniek-Klimczak, E. (Eds.), *Issues in teaching, learning and testing speaking in a second language*, 143-158, Springer.

To Cite this Chapter:

Önal, A. (2022). Teaching and assessing speaking. In A. Önal & K. Büyükkarcı (Eds.), *Essentials of foreign language teacher education*, (pp. 210-222). ISRES Publishing.

ABOUT THE AUTHOR



Dr. Ahmet ÖNAL

ORCID ID: 0000-0002-5325-4958

ahmetonal@sdu.edu.tr

Suleyman Demirel University

Dr. Ahmet Önal is currently working as an Assistant Professor Doctor in Süleyman Demirel University, Faculty of Education, Department of Foreign Language Education, Division of English Language Teaching. He received her Bachelor of Arts degree from Hacettepe University, English Language Teaching (ELT) department in 2003 and Master of Arts degree from Selçuk University ELT department in 2010. He completed his doctoral studies at the English Language Teaching Department, Hacettepe University in 2017. Önal offers courses at undergraduate and postgraduate degree. English language teacher education, teaching of language skills and technology-enhanced English language teaching & learning are among the fields of his interest.