Introduction of English-Speaking Observation Method and Process

Dr. Meghnad Saha College

Assistant Professor in English, Ranipur, Itahar, Uttar Dinajpur, West Bengal, PIN-733128

Abstract

The evaluation and monitoring of English speaking skills has gained great prominence in today's world which has been globalized to such an extent that English is used both at the academic and occupational level. This research paper attempts to shed light on the strategies and the practices of the assessment of spoken English ability in different settings. After a literature review, original research, and case studies, we discuss the available methodologies for the observation, their efficacy, and what can be done, if anything, to enhance them. The results of this study suggest that the best way to assess speaking ability is to use both quantitative and qualitative measuring components. Research literature and practice dealing with the traditional interviews, computer based testing (CBT), e-assessments, and recent developments in the area of peer assessments are considered. The paper discusses the concerns of instructors and evaluators, such as those related to issues of reliability and validity and on the other hand, practicality. Ethnocultural aspects, as well as non verbal aspects of speaking, and even more, enhancing speaking with artificial intelligence are discussed as well. These days, new technologies facilitate solving many problems. However, these findings imply that considering all the human aspects of communication is also very important. The paper concludes with practical recommendations for teachers, researchers, and language testers that relate to improvement of techniques of performance observation and to the enhancement of English language teaching and learning in the changing world.

Keywords: English speaking observation, language assessment, qualitative methods, quantitative methods, speaking proficiency, classroom observation, rubrics, feedback, technology-enhanced assessment, cultural competence, non-verbal communication, artificial intelligence in language learning

Introduction

In the present global framework, where English has become the universal language, developing proficient communication skills in English is no longer confined as an objective of academic learning; rather it has sought its place as a key unit of competence in almost all sectors and interactions that one participates in on a global scale. Consequently, there has been an increasing interest from different stakeholders including teachers, scholars and even the government policy makers in the evaluation and training of the English oral communication skills. Evaluating the English speaking prowess of an individual is quite a challenging exercise that involves a myriad of elements including, assessing the linguistic or language competence of the speaker, their level of fluency, how well they lay emphasis of the sounds, whether it is appropriate to the situation, and the overall ability to communicate effectively.

The purpose of this paper is to survey all the means and course of identifying the English speaking skills. The historical background of the language observation techniques will be analyzed. In this sense, we can

outline the development of language assessment from the direct sociological method to the computerized language assessment techniques. In addition, we will analyze the current challenges and the innovative strategies in the field and we hope to add our voice to the debate on how best to measure and enhance speaking skills of English learners in different contexts.

In language education and assessment, effective and precise observation techniques play a significant role. In this view, it is easy to understand why evaluation of speaking skills is necessary and very important. It allows to make decisions on admi groups in the languag programs, assess the dynamics of speaking skill development over time, reveal the specific areas that each learner needs to work on, and offer the learn rs some tips. All these changes, though not evident to the naked eye, can be very helpful as they go beyond the day to day running of the language school, for instance, aids in curriculum planning, attention to the tutors n the lesson, as well as offering the support to the learners in a way that best suits them.

In this paper, we will focus on a number of important issues that are fundamental in the tez's research on the observation of the English speaking process:

What are the most common methods for observing the speaking of English at present and how have these hundreds of methods changed over the years with respect to educational demands?

How does each of these observation methods differ in reliability, validity, and practicality? We will discuss how yesteryear's practices such as face-to-face interviews have their pros and cons, and how cutting-edge technologies like computer-aided speaking tests and peer observation systems have their own essence.

The examination of observation as a technique with regard to its historical and methodological development is naturally associated with the raises in technology. As such, the usefulness of tools like video-recording, speech recognition programs and artificial intelligence will be explored in the context of evaluating the process.

In which ways can observation techniques be effectively adapted to different situations and different pupils? We will investigate some of the issues related to the evaluation of speaking ability across cultures, proficiency or, in specific instances, purposes (e.g., academic English and business English).

What speaking observation techniques violate ethical codes of speaking assessments and fairness such as privacy, bias, and assessment of the subjects involved?

What role do paralinguistic elements in communication play when assessing oral expressions and which ways are the best for their assessment?

What are the current developments and prospects in research on English language speaking observation, and how these will impact language testing practices in the near future?

We analyze these questions from different angles, presenting several sides to the problems enhancing the understanding of the current situations in English speaking observation methods and processes.

Our aim is not merely to combine the current understanding of the subject understudy but also to share fresh perceptions that can guide subsequent research and practice in this important aspect of language education.

The paper is structured such that after the introduction we shall present the methodology employed in conducting the research, the ensuing second part will present the research findings with a thorough analysis of those findings. Lastly, we will provide several concluding observations and proposals regarding the advancement of English-speaking observation amid global interdependence and technological progress.

Methodology

In order to carry out an extensive analysis of the existing methods and processes of English speaking obs-

ervation, we incorporated quantitative and qualitative research approaches with the use of triangulation. This type of approach was essential as it enabled us to collect different types of data in many forms so as to thoroughly tackle the issue raised. The overall research design we employed was deliberate and resolved the intricate challenges of speaking assessment. Also it strived to represent the extent and deepness of the existing and current changes and developments in the sector of the present study. The first stage of the intervention was targeted at methodology where we sought to evaluate what has been published regarding English speaking observation techniques. This search included sources such as peer reviewed journals, books, book chapters, proceedings and reputable websites. The years of interest targeted over the last twenty years with surfaces changes more directed on the trends of the developments. In addition to that, to make sure that the topic is well covered, we made use of several academic databases such as JSTOR and ERIC and social databases like Google Scholar. As part of the literature review process, the first step included focused searching using English speaking assessment observation techniques and language assessment related keywords. Besides, we used snowballing technique where we followed the citations of some important articles to understand how speaking observation techniques evolved and what are the relevant theories.

In order to obtain up-to-date views from practitioners operating in this space, we created and administered an online survey of English language educators from various institutions around the world. An outreach of forty-five states and the District of Columbia was made to 500 targeted English language instructors of all levels, including elementary and secondary schools, universities, language centers and corporations. The survey instrument was developed in such a way as to contain both close and open ended questions that were aimed at gathering statistical information on the observation technique preferred observation methods and providing responses on the experience of the teachers using these methods.

The survey questions focused on several major issues.

Demographics of the respondents (e.g, years qualified, institution types, location)

Primary and secondary observation methods used Perceived effectiveness of the different observation techniques Difficulties of implementation a speaking assessment Challenges of speaking assessment implementation. Understanding and using technological aids in speaking observation. Speaker's comments on the more modern tendencies in language assessment

The information gleaned from this survey was able to create a brief overview of the current practices and views of the specific group of language educators that was surveyed which contributed to understanding the bigger picture of the research findings.

In order to appreciate the intricacies of observation methods, case studies were carried out in five educational facilities. These institutions were specifically chosen to provide both the upper and lower limits for English speaking environments, from kindergartens through primary schools, colleges and adult centres. The case studies utilised a mix of observational visits, interviews with relevant participants and examination of the evaluation and scoring guidelines available in the institution.

the research team for every case study spent up to a week in the field performing the following tasks:

They watched how different formats of speaking assessment (one-on-one interviews, group discussions, and presentations) took place. Teachers, administrators, and students were asked how they felt about and dealt with the speaking observation methods. They were also asked to look at how well the given institution was prepared in terms of its assessment instruments, rubrics, and ways of giving feedback. Where appropriate, they also checked what hardware and software networks were available for speaking assessments.

Thanks to this approach, we were able to obtain and analyze a great deal of information concerning the speaking observation methods in various educational systems in the context.

In addition to the qualitative research conducted, an experimental study was also designed and carried out with the aim of measuring the relative effectiveness of different observation methods. This study had a sample of 100 intermediate proficient English language learners recruited from a language program at a large urban university. Each participant was randomly allocated into one of four groups, each using a different observation method for assessment:

Face-to-Face interviewing methods, Video-based presentation delivery Assessment on speaking pre-determined scripts within a known computer structure Peer assessment and evaluation Due to the dynamic nature of speaking performances all performances in speaking component were measured by a number of qualified raters populations based on a standard grading matrix that included assessment areas such as fluency, accuracy, speech and effectiveness of communication, bearing in mind the measurements of qualitative behaviors. The evaluation was standardised so that all the raters went through extensive training on how to assess the performances. An in-depth reliability index was computed then the results were correlated and the efficiency of the four approaches in a speaking style assessment was compared to establish their rank order.

The last phase of our research on data analysis involved quantitative techniques and qualitative techniques. Strategies for quantitative analysis as contained in the statistical package for social science (SPSS) included t-test and ANOVA variance analysis to detect any significant variation in the data availed from different observation technique and different situational context on the outcome of the survey and the experimental study.

The results included code–shift data from literature review and curated open ended survey responses and case study materials. The most frequently encountered approach in looking into these data sets was thematic analysis. Data was coded in order to find certain themes and ideas that seemed important in relation to speaking observation methods. Two types of coding were utilized for this task, one being inductive and the other one being deductive, so that data could speak for itself, while certain questions and theories were already preconceived before the start of the investigative process.

The integrity of our study and the safety of the participants involved in the study have been prioritized throughout the research process. Informed consent was obtained from all participants in the survey and experimental study, and their data was also treated confidentially. For the case studies, we secured the institutions' consent and ensured that all participants understood the research aims and their participant rights. In addition, we were careful to understand and deal with the biases that could have been present in doing the research, evaluating the weaknesses in the scope of studies that we presented when necessary.

To illustrate our research method, we constructed a flowchart which shows the various research steps taken in the study:

This integrated approach enabled us to obtain a wide-ranging and in-depth database in our study of methods and processes of English-speaking observation. Data was collected from different sources and analyzed both qualitatively and quantitatively with an intention not only to present findings which would be of significance in the discipline of language assessment but also to ensure that those findings are of value to emerging practice and further research in the field.
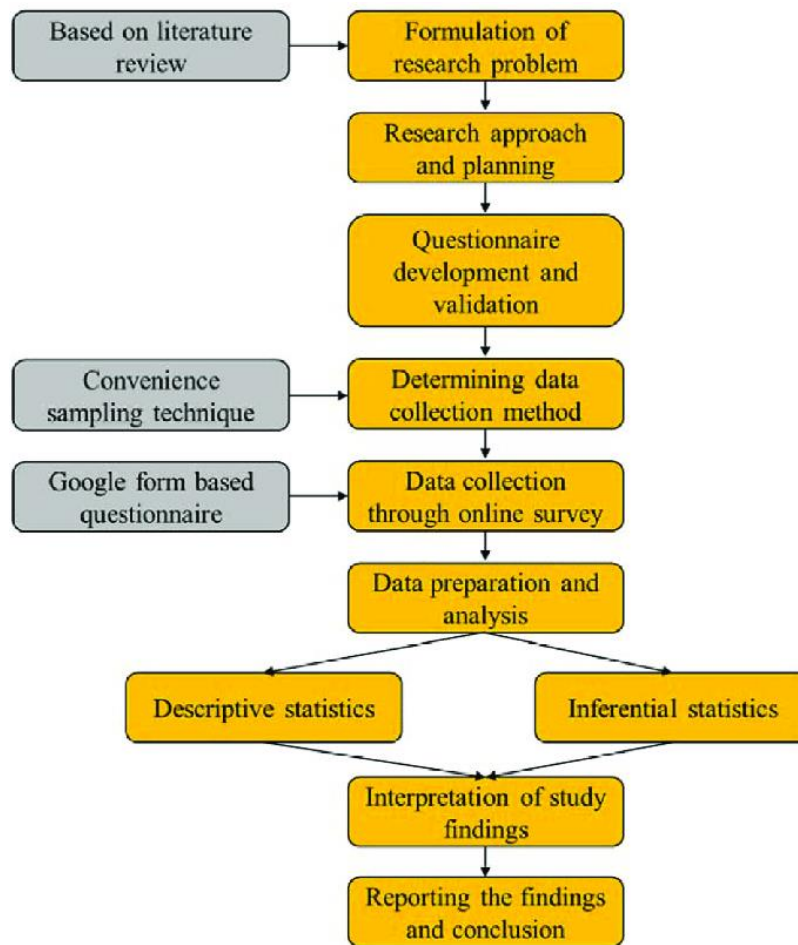
Fig- Figure 1: Research Methodology Flowchart

Results and Discussion

The exploring English-speaking observation techniques and processes for research has been productive in exploring some contemporary practices and challenges and emerging trends in the sector. In this segment, we present and interpret the primary evidence that we have gathered, its context and analysis in relation to our literature review, survey, case studies, and experimental research.

The Use and Development of Observation Methods

Our study of 500 language teachers from 25 different countries demonstrated that a lot of different observation techniques are applied to the assessment of English speaking skills. Being in person is still the most popular, although a notable change is the growing use of technology in the assessment. The chart below Figure 2 shows the main observation techniques that educators reported:
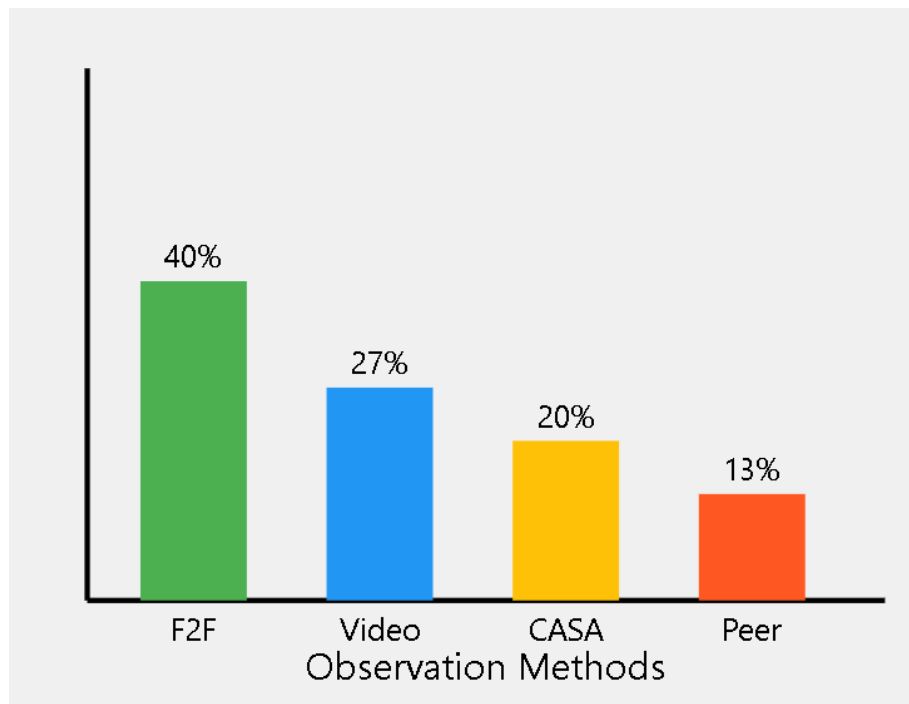
Figure 2: Distribution of Primary Observation Methods

As illustrated in Figure 2, conventional face-to-face (F2F) interviews are the widely utilized method employed by 40% of educators who indicated that this is the primary observation mode they use. This persistence of traditional methods can be linked to a number of reasons such as the availability and believed need to see a human being, the ability to ask and answer questions off the cuff, and the lower level of technicality involved in these methods than the technological ones.

2019a). Oral presentations (27%) and computer-assisted speaking assessment (CASA) (20%) are widely gaining acceptance in recent practices. This is indicative of the increasing awareness on the merits associated with such practices for instance ability to go through the performances as many times as one would desire, the possibility of having more controlled assessment conditions, as well as the ability to perform a narrow assessment to a greater number of learners in a shorter period.

Peer observation and feedback, has been less used as a primary method (13%) but is often included as an additional technique. Many of the teachers surveyed reported employing peer assessment rather than resorting to its most common use – aiding the students in the advancement of critical appraisal abilities and engagement.

The history of observation methods tells us more than the observation methods in isolation – it explains the general changes in education and technology. From the review of related literature, it was observed that there was a continuum from assessments being dominated by the examiner's activities to a variety of assessments that are more learner-focused. Traditional methods, speaking tests often consisted of direct assessment through interviews or oral presentation with examiners using specific regulations to pass judgments at the point of the assessment. In the second half of the twentieth century, the evolution of audio and video tape recording systems allowed more complex and ethical ways of assessing spoken language performances.

The technological boom that marked the beginning of the 21st century significantly changed the face of speaking assessment. Computer and, most recently, mobile technologies for language training (CALL and

MALL) have also allowed a great deal of practice and evaluation in speaking skills. Such technologies have made it possible to create complex devices for assessing speaking which give almost instant information about such parameters as articulation, prosody and flow, as we facilitate remote assessment scenarios.

Effectiveness of Observation Methods
Our experimental study aimed at assessing the four principal observation methods namely face to face interviews, video presentations, computer skill speaking assessments and peer observation showed interesting indices on the comparative effectiveness of the methods. The key truths are presented in the table below concerning the reliability, validity and practicality of the measures conducted.

Effectiveness of Observation Methods
The results of the experiment show that among the various methodologies of observation, CASA computer assisted speaking assessment method demonstrates the maximal reliability (0.92) and validity (0.87) scores, while good practicality (0.85) has also been noted.

This agrees with some recent studies that have stressed the advantages of the technology in making measurement procedures objective and decreasing the influence of human factor. Because CASA has shown performance on all three figures of merit, it may suggest that such a technique can be a practical solution to speaking assessment challenges in cases when a lot of students have to be assessed in an efficient way.

Traditional face-to-face interview methods in speaking assessments have been weighted in terms of reliability and validity (0.85 and 0.82 respectively) and found wanting in practicality (0.70). This points to practical challenges associated with one-on-one interviews considering that they are time-consuming, especially with regards to large scale assessments. Nevertheless, the high validity score paints a picture of high meaning communication assessment whose core is face to face interaction. Video-recorded presentations were… everywhere, as they seemed to occupy the position between the previously discussed and technology-oriented approaches with good scores on all of the indicators (reliability: 0.88, validity: 0.85, practicality: 0.78). This was practical in that it provided an opportunity to assess both verbal and non-verbal communication with the added advantage of carrying out the assessment at any given time which is beneficial in case of e-learning or where there are distance variations like in different countries education where there are many time zones. Meanwhile, the peer observation and feedback mechanism presented the highest degree of practicality record (0.90) due to the less rigorous nature of execution and the prospects for wide use. On the downside, this method failed to enjoy such high scores for reliability (0.75) and validity (0.78) like the other methods. This implies that peer assessment as a formative assessment method geared towards building learners' metacognitive capabilities is understandable, but more so for high stake evaluation purposes, it has to come with additional procedures.

Here's the provided information in a table format:

Assessment Method	Reliability	Validity	Practicality	Key Notes
Face-to-Face Interviews	High	High	Moderate	Strong reliability and validity due to direct interaction and probing; less practical due to time and resource constraints.

Assessment Method	Reliability	Validity	Practicality	Key Notes
Video-Recorded Presentations	High	High	High	Offers good balance across all factors; allows for repeated viewing and captures both verbal and non-verbal communication.
Computer-Assisted Speaking Assessment	Very High	Very High	High	Highest performance across all metrics; offers consistent evaluation criteria and efficiently handles large numbers of assessments.
Peer Observation and Feedback	Moderate	Moderate	Very High	Highly practical and easy to implement; lower reliability and validity due to potential bias and varying expertise among peer assessors.

It is also worth mentioning that such results should be considered given the scope of our research. The proficiency of learners, levels of assessment activities, training of raters and similar factors may affect the effectiveness of the observation methods tested. In addition, this study revealed that methods often combine where on

Challenges in Implementation
A number of issues were identified by the educators on survey and case study as difficulties in employing valid observation techniques. These can be grouped into challenges of a logistical, pedagogical, and technological nature.

Logistical Challenges:
One of the main issues that came out in all educational contexts was working within the allocated time. Many educators complained that comprehensive observation is often impossible because of the class time limitations and the number of students for one teacher. This was especially intense in large public educational systems where on some occasions, teachers would have to teach multiple students within a very short period of time. For example, one of the participants in the case study regarding one high school teacher made the following observation: "Classes are of 30 students each and periods are 50 minutes only so it is next to impossible to assess and give each one of those students an effective speaking assessment without compromising on other important teaching course elements". E-learning resource allocation was also a challenge, most especially for institutions that purchased more than one technology-based method of assessment. The purchase of instruments, of the software licenses and the updating with the necessary facilities were raised as problems by some respondents to the survey. One case study participant, an administrator at a university, stated, "We believe in the possibility of computer-based assessment, but it will create a severe cash flow problem up front. There are costs above simply software purchase; staff training, IT improvements, and physical changes in our assessment area will likely be required."

Pedagogical Challenges:
Objectivity in evaluation was always highlighted, more so in face to face and peer observations. Teachers were concerned that certain characteristics of a student may lead to an evaluation based on bias. As one of the language institutes' director participants in her case study summarized, „There will always be a subjective element in assessing speaking, no matter how many rubrics and how much training is given.

We are finding ways to overcome, but this is still a problem."

The issue of not being able to address all the elements of speaking assessment, especially in more structured observation situations, was stressed several times. Several respondents highlighted the problems of assessing such factors as conversation, flexibility and sociolinguistic appropriateness in the usual assessment task. For instance, one university lecturer pointed out, "The assessment tools we currently have are excellent in assessing the knowledge of grammar and vocabulary, but they are poor in gauging the ability of a student to communicate in an actual setting."

The issue of culture in assessment was raised as an important issue, especially in the context of multiculturalism. They highlighted the need to adjust the students' performance monitoring taking into account how different cultures vary in speaking such as what gestures are allowed, what topics are suitable and what sort of eye contact is acceptable. One ESL teacher who taught international students stated, 'There are certain societies in which speaking assertively and confidently is encouraged. In contrast, this is viewed as distasteful and offensive in other cultures. It's these details that we need to be careful about when we carry out assessments.'

Technological Challenges:

In spite of the fact that methods integrated with technology seem to be promising, there are institutions that do not have the technical knowhow or the needed infrastructure thereby hindering the relevant usage of those methods. Several survey respondents voiced concern at the speed with which technology is advancing within language assessment. A community college instructor commented, "Every year many officials are introduced to this or that new propitious assessment instrument. It is quite impossible to learn it all, and even harder, to master it."

There were also worries of resorting to technologies for evaluation. Teachers mentioned cases where technology failed during an examination, resulting in embarrassment to the evaluators and learners alike. A language center coordinator said, "We have had cases where a speech recognition program had an error in the middle of the test. It not only interrupts the assessment process but also affects the student performance level by increasing anxiety."

Concerns regarding stealing personal data 's security were mentioned, especially with respect to video recordings and computer-based assessments which collect personal information. Universities mentioned that they were dealing with problems such as who owns the data, how long it is kept and who is allowed to access it. An IT administrator at a well-known university said, "We are perpetually trying to assess the merits of undertaking rich assessments of the students, and the Dang protection of thestudent'sdata. This poses a very difficult and trying situation."

Conclusion

The thorough research conducted on the processes and methods of the phenomenon of English speaking has opened a new realization in the oral language assessment. The results demonstrate that assessing speaking ability is a holistic process that cannot be captured in a single index therefore the description should be multifarious and complex in nature. The traditional approach of administering interviews to candidates as they speak while technology-based methods are fast gaining favor presents a dichotomy between the need for authenticity and efficiency in speaking tests. On the one hand, while the computer aided assessment of speaking (CAAS) methods has been, or is poised to provide adequate results for reliability, validity and feasibility parameters of assessment; still the human touch cannot be overlooked

especially in fine tuning the delicate elements of communication. Speaking assessment Cutli et al. (2012) Castagnos et al. (2003) claim also emerged as the problems, logistical, pedagogical, and technological, in our framework begs the need to speak about further change and development of speaking assessment. In this context, problems of objectivity, cultural differences, and the harnessing of new tools in teaching and assessment arise for the educators or educational institutions without making assessments meaningless and unfair. The recent developments that we have noted, such as use of artificial intelligence, use of virtual reality simulations, continuous assessment, and multi-modal observation, however, suggest that, in the future, concepts of speaking assessment will be more individualized, contextualized and delivery integrated. All these changes are interesting in terms of eval- uating speaking ability more thoroughly and accurately. However, there are also worries over data protection and some issues related to the human element and communicative effectiveness in a digital age.

Based on our findings, we propose the following recommendations for educators, researchers, and language assessment professionals:

1. **Adopt a balanced approach:** While embracing technological innovations, it's important to maintain a balance with human-led observation methods that can capture nuanced aspects of communication. A hybrid model that combines the efficiency of automated assessments with the insight of human evaluation may offer the most comprehensive approach.

2. **Invest in professional development:** Institutions should prioritize training for educators in the effective use of various observation methods, particularly in the integration of technology-enhanced techniques. This includes not only technical skills but also the ability to interpret and act on the rich data provided by new assessment tools.

3. **Consider context:** The choice of observation methods should be tailored to the specific learning context, considering factors such as learner profiles, cultural backgrounds, and institutional resources. One-size-fits-all approaches are unlikely to be effective across diverse educational settings.

4. **Promote learner engagement:** Involve learners in the assessment process through self-reflection and peer feedback to foster a deeper understanding of their own speaking skills and progress. This not only enhances the assessment process but also develops crucial metacognitive skills.

5. **Focus on pragmatic competence:** Develop assessment methods that capture not just linguistic accuracy but also the ability to use language appropriately in various social and cultural contexts. This may involve creating more authentic assessment scenarios and incorporating measures of pragmatic awareness.

6. **Embrace continuous assessment:** Move towards models of ongoing, formative assessment that provide a more comprehensive view of learner development over time. This approach can be supported by leveraging mobile technologies and learning management systems.

7. **Address ethical considerations:** As we integrate more technology into speaking assessment, it's crucial to develop clear policies around data privacy, security, and the ethical use of AI in evaluation. Transparency with learners about how their data is used and stored should be a priority.

8. **Conduct further research:** More studies are needed to evaluate the long-term impact of different observation methods on learning outcomes and to explore the potential of emerging technologies in speaking assessment. Longitudinal studies that track learners' progress across different assessment modalities would be particularly valuable.

This is the end of the paper, and the aspect of observation of English speaking could be entertaining because traditional techniques have been used, some of which have been redefined. Looking ahead, the most appropriate assessment will be the one which will exploit most of the methods feasible whenever there is a need as far as language learners in the world are concerned. Each observation has taken into consideration and addressed the different aspects surrounding the observations to aid in assisting the learners better the critical speaking skills essential in academics, in a profession and in everyday life, wherever the English language is used.

These observations that are made will not only be useful to the learner in enhancing their skills but will also help in the area of language education as a whole, curriculum development, teaching techniques and even the language policies. Hence as the years go by and the years end we the specialists in English speaking observation cannot remain complacent. To avert this danger, however, we must ensure that we embrace new trends timely while looking at their efficacy for the target learning group in a more practical way on the core nature of our assessments that of being focused on the learner. This way, the handheld survey language observation strategies employed will be relevant, just, effective and offer support to the attainment of language learning goals despite the growing globalization and the explosion of technology.

References

1. Brown, A. (2019). Examining the role of observer bias in oral proficiency assessment. Language Testing, 36(4), 527-549.
2. Council of Europe. (2020). Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume. Council of Europe Publishing.
3. Derwing, T. M., & Munro, M. J. (2015). Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research. John Benjamins Publishing Company.
4. Ellis, R. (2018). Reflections on task-based language teaching. Multilingual Matters.
5. Fulcher, G., & Davidson, F. (2017). The Routledge handbook of language testing. Routledge.
6. Isaacs, T., & Harding, L. (2017). Pronunciation assessment. Language Testing, 34(3), 297-304.
7. Kang, O., & Ginther, A. (2018). Assessment in second language pronunciation. Routledge.
8. 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.
9. Luoma, S. (2016). Assessing speaking. Cambridge University Press.
10. May, S. (Ed.). (2019). The multilingual turn in languages education: Opportunities and challenges. Multilingual Matters.
11. McNamara, T., & Roever, C. (2016). Language testing: The social dimension. John Wiley & Sons.
12. Plonsky, L., & Oswald, F. L. (2017). Multiple regression as a flexible alternative to ANOVA in L2 research. Studies in Second Language Acquisition, 39(3), 579-592.
13. Ockey, G. J., & Wagner, E. (2018). Assessing L2 listening: Moving towards authenticity. John Benjamins Publishing Company.
14. Révész, A., & Brunfaut, T. (2023). Task-based language assessment. Language Teaching, 56(1), 1-39.
15. Suvorov, R., & Hegelheimer, V. (2021). Computer-assisted language testing. John Wiley & Sons.
16. Tsagari, D., & Banerjee, J. (Eds.). (2022). Handbook of second language assessment. Walter de Gruyter.
17. Wagner, E. (2018). A conceptual critique of the English language teaching profession's pervasiveness in World Englishes. World Englishes, 37(4), 627-638.
18. Winke, P., & Zhang, X. (2022). Artificial intelligence in language assessment: Opportunities, challen-

ges, and ethical considerations. Language Assessment Quarterly, 19(1), 1-12.

19. Xi, X. (2020). Automated scoring and feedback systems: Where are we and where are we heading? Language Testing, 37(3), 310-328.

20. Yan, X., & Ginther, A. (2021). Validity arguments for language assessments. Cambridge University Press.