

الاختبارات الشفوية في تعليم اللغة الثانية

دراسة في ضوء علم اللغة التطبيقي^(*)

د. غسان حسن الشاطر

أستاذ مساعد - قسم اللغة العربية وآدابها -

كلية الآداب - جامعة اليرموك، إربد، الأردن.

الملخص:

تسلط الدراسة الحالية الضوء على الاختبارات الشفوية في تعليم اللغات، بما في ذلك استخدام الحاسب ((Computer Assisted Language Learning (CALL) في الاختبارات الشفوية للغات الثانية. وتركز على مفهوم مهارة المحادثة، وأساليب تقييمها، وفوائد التقييم، ومواصفات الاختبارات، وأبرز التحديات عند اختبار المحادثة، ووضع أسس معيارية يمكن اعتبارها متطلبات رئيسية لتطوير برامج حاسوبية تساهم في التقييم الآلي لمهارة المحادثة. كما تهدف إلى وضع مواصفات معيارية تستخدم في تصميم أداة لقياس المهارات الشفوية من جهة، وتشكل خطوة في طريق التقييم الآلي من جهة أخرى. ولتحقيق هذه الأهداف فإن الدراسة الحالية ستقدم عرضاً وصفاً لعدد من الدراسات السابقة التي تناقش مفهوم مهارة المحادثة، وتقويمها، ومواصفات مقياس التقييم، وتحديات اختبار المحادثة، ثم تبين فوائد تطوير برنامج خاص بتقييم المهارات الشفوية لدى متعلمي اللغات الثانية، ومواصفات تلك الاختبارات، ويعد ذلك تقترح أبرز الإجراءات اللازمة لتطوير هذا البرنامج.

تظهر نتائج الدراسة أن هناك حاجة ماسة إلى تطبيق معايير تقييم كمية لقياس أداء متعلمي اللغة الثانية في مهارة المحادثة، وضرورة العمل على إيجاد آلية لتحويل المعايير النوعية إلى كمية، كما تشير إلى ضرورة وضع جداول مواصفات لتلك المعايير قبل استخدامها في تطوير برامج التقييم الآلي؛ لضمان المصداقية، والعدالة، وسهولة عقد الاختبارات الشفوية.

تخلص الدراسة إلى نتيجة مفادها أن تقييم مهارة المحادثة ضروري لضمان اكتساب اللغة الثانية، واستخدامها في التواصل، وتوضح أن تطوير برامج آلية لتقييم مهارة المحادثة لدى متعلمي اللغة الثانية ممكن، وأن مثل هذه البرامج سيساعد في ضمان أكبر قدر من المصداقية، والعدالة، ويقلل من التعقيدات التي يواجهها مقيمو مهارة المحادثة.

الكلمات المفتاحية: اختبارات، تقييم، لغة ثانية، محادثة، تعليم.

Oral Tests in Second Language Teaching An Analysis Inspired by Applied Linguistics

ABSTRACT

This study addresses a critical aspect of language education: the oral testing of second languages, including the integration of computer technology into language teaching. By illuminating the concept of speaking skills, various assessment methods, and the benefits of effective evaluation, this research aims to establish clear standards essential for developing innovative computer programs designed to automate the assessment of speaking skills.

To fulfill these goals, the study will present a thorough overview of existing research on conversational skills and their evaluation, including the specifications of assessment scales and the common challenges faced in conversation testing. It will underscore the necessity of creating a tailored program to assess the oral competencies of second language learners and will detail the optimal characteristics of such assessments.

The findings highlight the urgent need to implement quantitative evaluation standards that accurately measure the speaking abilities of second language learners. Developing automated assessment programs is crucial to ensuring that oral exams are conducted with integrity, fairness, and efficiency.

Ultimately, this study underscores the importance of conversational skill assessment in facilitating second language acquisition and effective communication. It presents a compelling case for developing automated programs that enhance reliability and fairness in evaluating conversational skills while alleviating the complexities encountered by assessors.

Keywords: Tests, Evaluation, Second Language, Speaking, Teaching.

مقدمة

تعتبر اختبارات قياس قدرات التحدث لدى المتعلمين من المهام المعقدة في تدريس اللغات بشكل عام (Sandlund & Sundqvist, 2021, p. 132; Köroğlu, 2019, p. 23)، إلا أن المهمة تصبح أكثر تعقيدا عند الحديث عن تدريس اللغات الثانية؛ وذلك لعدة أسباب، من أبرزها: أن عددا كبيرا من المؤسسات المعنية بتدريس اللغات الثانية تركز على تدريس مهارتي القراءة والكتابة، وتدرس العناصر اللغوية (النحو، الصرف، المفردات)، ولا تعير الاهتمام نفسه لتدريس المحادثة (Torky, 2006, p. 21)، بل إن غالبية المدرسين يتجاهلون تدريسها (Köroğlu, 2019, p. 29) مما يجعل العمل على اختبار مهارة المحادثة قليلا، ويقتصر على أدنى متطلبات تقييمها، وربما يخلو نظام تدريس اللغة الثانية أحيانا من "أي تقويم يتعلق بقياس مهارة المحادثة، وهي مشكلة تتطلب عناية رسمية وعاجلة" (جيدور، ٢٠٢١، صفحة ١٥٦)، ولا تعد العربية استثناء (الشاطر، ٢٠١٩).

ويلاحظ أن الاهتمام بتدريس مهارة المحادثة ازداد في العقدين الأخيرين على مستوى عالمي، وفي العقد الأخير على مستوى تعليم العربية لغة ثانية، فيما يمكن اعتباره تقدما ملحوظا في النظرة إلى تدريس المحادثة، أدت إلى ازدياد الاهتمام بتصميم أنشطة المحادثة، مثل: المقابلات، ولعب الأدوار، والمجموعات التعاونية، والخطب، والنقاش، والحوار، وحل المشاكل، ووصف الصور، والصف النادي، والتفاوض، وغيرها (Iwashita, Lyn, & Moore, 2021, p. 287; Kuśnierek, 2015, p. 80; Arung, 2016, p. 72; Pratiwi & Ayu, 2020, p. 40).

أدى ازدياد الاهتمام بمهارة المحادثة في تدريس اللغات إلى ازدياد الحاجة إلى تقييم أداء الطلبة في هذه المهارة، الأمر الذي استدعى التفكير في طرق مناسبة

للتقييم تضمن المصادقية والثبات، ويمكن الاعتماد عليها، في تحديد مستويات الطلبة (Carrió-Pastor, 2016, p. 62; Al shatter, 2017; Suvorov & Hegelheimer, 2014, p. 11). ولعل من هذه الطرق ما يسمى بسياسة التوفيق؛ وفيها يطور أعضاء فريق التدريس فهماً مشتركاً (توافقياً) لمتطلبات التقييم والمعايير والأدلة التي توضح الصفات المختلفة للأداء، قبل الشروع في تقييم الطلبة، وهو نشاط تقييمي جماعي يشترك فيه المقيّمون، أو يتعاونون على تقييم أداء المقيّمين، أو يناقشون التقييمات الفردية لأدائهم (Sandlund & Sundqvist, 2021, p. 133).

وتشير الدراسات إلى عدد من طرق التقييم التقليدية، مثل: قوائم التقييم (Chick Lists)، وقواعد التقييم (Rubrics)، التي تسعى إلى تقييم متحدث اللغة، وتزويده بتغذية راجعة تفيد في تطوير قدراته التواصلية عبر التحدث (Köröğlu, 2019, p. 24). إضافة إلى الطرق الحديثة، مثل: التدريس القائم على الذكاءات المتعددة (Multiple Intelligences)، الذي يعتمد على ثماني قدرات أساسية يحتاجها المتعلم لتعلم المحادثة (موسيقية؛ الأصوات بشكل عام، بصرية، لغوية؛ بما في ذلك المفردات، المنطق الرياضي؛ بما في ذلك التفكير الناقد والأرقام، القدرات الجسدية؛ بما في ذلك القدرة على تقدير الوقت والتحكم بالأفعال، القدرات البدنية؛ بما في ذلك التحكم بأعضاء الجسم، قدرات شخصية، قدرات شخصية تفاعلية؛ كيفية التحكم في التعامل مع الآخرين). هذه العناصر الثمانية تساعد المعلم على الاستفادة من قدرات المتعلمين المختلفة للتدريب على المحادثة وممارستها (Rizqiningsih & Hadi, 2019, p. 131; Altun & Sabah, 2020, pp. 149-155).

من هنا، غدا التفكير في حوسبة اختبارات المحادثة للغات الثانية متوقعا؛ حيث أتاحت برامج التعرف الآلي على الحديث (Automatic Speech

(Recognition (ASR) للغويين فرصاً مكنتهم من تحديد هوية المتحدثين بناءً على النطق، والحكم على جودة النطق لديهم وبدرجة عالية من الدقة مقارنة بتقييم البشر، وتحديد أنواع الأخطاء في النطق، واكتشاف أخطاء الطلاقة اللغوية، وتزويد المتعلمين بالتغذية الراجعة (Carrió-Pastor, 2016)، مما يشرع الآمال في أن تتمكن هذه التقنية من التعرف إلى حديث متعلمي اللغات الثانية وتقييمه.

ضمن هذا السياق تكونت الدراسة الحالية من عرض لأبرز القضايا المرتبطة باختبارات المحادثة، بدءاً بتعريف مهارة المحادثة، وطبيعتها، وعرض أبرز الآراء التي تناقش تصنيفها، ومفهومها، ودور المعلم في تعليمها، ودور المتعلم في تعلمها، مروراً بالممارسات المهنية في تقييمها، وأركان التقييم التي تشمل المقيّم، والمقيّم، ومقياس التقييم، ثم توضيح الفوائد التي تعود على المتعلمين جراء تقييم مهارة المحادثة، ومواصفات اختبارات المحادثة، وأبرز التحديات التي تواجه الخاضعين لتلك الاختبارات، ومصممها، مع عرض لأنواعها، مع التركيز على اختبارات المحادثة المحوسبة، والدور الذي يمكن أن تلعبه في مواجهة تحديات تصميم اختبارات المحادثة وعقدها، ثم تنتهي الدراسة الحالية إلى خلاصة تبين أهمية تدريس المحادثة، وارتباط نجاح العملية التدريسية بتقييم قدرات المتعلمين التفاعلية والتواصلية، في أثناء التحدث باللغة الثانية.

مشكلة البحث:

تعالج الدراسة الحالية الاختبارات الشفوية لمهارة المحادثة في تعليم اللغات؛ حيث تناقش الممارسات الراهنة، وما تحتاج إليه من عمليات تطوير، منطلقة من مفاهيم رئيسة في فهم أهمية اختبارات المحادثة، وطبيعتها، وأركانها، والتحديات التي تواجه المدرسين والمتعلمين على حد سواء، محاولة التوصل إلى حلول تعالج تلك التحديات، وخطوات التغلب عليها، بما في ذلك الاختبارات الآلية؛ من أجل النهوض بمستوى متعلمي اللغة الثانية في مهارة المحادثة، وتصميم اختبارات مناسبة للمقيّمين، والمقيّمين.

أسئلة البحث:

تهدف الدراسة الحالية إلى الإجابة على الأسئلة الآتية:

١. ما الدور الذي تلعبه اختبارات المحادثة في تنمية قدرات متعلمي اللغة الثانية في مهارة المحادثة؟
٢. ما هي أبرز أشكال اختبارات المحادثة، والمواصفات المرتبطة بكل شكل منها؟
٣. ما هي أبرز التحديات التي تواجه مصممي اختبارات المحادثة، وأنجع السبل في تجاوزها.

منهج البحث:

تعتمد الدراسة الحالية المنهج الوصفي التحليلي في تتبع الممارسات الراهنة في تدريس المحادثة واختبارها، مع التركيز على اختبارات المحادثة، وعرض أبرز ما يتعلق بها من قضايا تساعد في فهم أهميتها، والتحديات التي تواجه المهتمين بها؛ من مدرسين، وطلبة، وسبل تطوير تلك الاختبارات؛ للحد من تلك التحديات، ورفع قدرات المتعلمين.

مهارة المحادثة:

يرى نزارا (2011, p. 30) أنه من الصعب التوصل إلى تعريف موجز وشامل للتحدث، وخير ما يمكن أن نحدده هو تعريف الكلام بخاصيته ووظائفه وظروفه، وتظهر التعريفات المختلفة لمهارة المحادثة عددا من العناصر التي تلعب دورا ضروريا في تعريفها، أبرزها ما أشار إليه بيرنز بأن التحدث عملية تفاعلية لبناء المعنى، تتضمن إنتاج المعلومات وتلقيها ومعالجتها. ويعتمد شكلها ومعناها على السياق الذي تحدث فيه، والمشاركين، وأغراض التحدث (Torky, 2006, p. 28; Rizqiningsih & Hadi, 2019, p. 129).

بهذا تكون المحادثة نشاطا لغويا يسعى فيه الإنسان إلى نقل أفكاره ومشاعره

ورغباته إلى الآخرين عبر مواقف تواصلية تفاعلية متعددة الأغراض والأشكال، بلغة سليمة، قادرة على عكس ما يفكر به ذلك الشخص بدقة ووضوح (Altun & Sabah, 2020, p. 148).

يعد التحدث بطلاقة في اللغة الأولى أو اللغة الثانية قدرة معرفية معقدة تتطلب تحويل المعرفة المباشرة الآنية إلى معرفة إجرائية وتلقائية (Mirzaei & Heidari, 2012, p. 35). وتصنف مهارة المحادثة على أنها من المهارات المنتجة (Productive) مع الكتابة، في مقابل مهارات الاستقبال (Receptive)، وهما: الاستماع، والقراءة، كما يصنفها عدد من الباحثين على أنها أكثر المهارات الأربع أهمية، وأكثرها صعوبة (Zakaria, Hashim, & Yunus, 2019, p. 3082)؛ لأنها تتطلب قدراً أكبر من الممارسة في بيئة مُتحدثة للغة المتعلمة، إضافة إلى أن كثيراً من متعلمي اللغات يعدون أنفسهم مجيدين للغة المستهدفة إذا تمكنوا من التحدث إلى ناطقي تلك اللغة الأصليين بنجاح (Kuśnierek, 2015, p. 75; Arung, 2016, p. 71).

إن الهدف الأبرز من التواصل هو تبادل المعلومات في سياق تفاعلي، ولا يعد هذا التبادل ناجحاً إلا إذا كانت المعلومة التي وصلت المتلقي هي نفسها التي أرسلها المتكلم (Kuśnierek, 2015, p. 77; Koroğlu, 2019, p. 23)، ولا يمكن أن تكون المعلومة واضحة، إذا لم يتمكن المتكلم من ترجمتها نطقاً بالشكل المطابق تماماً لما في ذهنه، حينها لن يتمكن المتلقي من فهم المقصود، وربما يأخذ الحوار منحى مختلفاً، أو يتوقف؛ مما يشير إلى محاولة فاشلة للتواصل.

تتمثل المهمة الرئيسة للمعلم في مساعدة المتعلم على إتقان العناصر اللغوية الأساسية (النطق والقواعد والمفردات)، وعناصر وظائف اللغة، وعناصر المعايير الاجتماعية والثقافية؛ من خلال توفير تمارين أصيلة تعدهم للتفاعل في مواقف التواصل الحقيقية. ويجب على المدرسين مساعدة المتعلمين على إنتاج التعابير

النحوية والمنطقية، وفقاً لحالة التحدث، بمستوى نطق مقبول (Nazara, 2011, p. 33). إضافة إلى ما سبق، فإن على المعلم أن يخلق فرصاً ومواقف تعليمية طبيعية تتيح للمتعلم ممارسة المحادثة في غرفة الصف عبر أنشطة محادثة مصممة خصيصاً لتحقيق هذا الهدف (Rizqiningsih & Hadi, 2019, p. 129). لعل من أبرز العناصر التي تظهر أهمية تعليم مهارة المحادثة هو عنصر التواصل الاجتماعي؛ فقدرة متعلم لغة ما - سواء لغة أولى كانت أم ثانية - على التحدث بها، تتحكم في مستوى قدرته على التواصل مع متحدثي تلك اللغة؛ فإتقان التحدث يزيد من مستوى التواصل، الأمر الذي يزيد بدوره فرص الاندماج في المجتمع الذي يتحدث اللغة المستهدفة (Kuśnierek, 2015, p. 76; Zakaria, 2015, p. 8035; Rizqiningsih & Hadi, 2019, p. 129). Hashim, & Yunus, 2019, p. 8035; Rizqiningsih & Hadi, 2019, p. 129).

تقييم مهارة المحادثة:

تنوعت أشكال وإجراءات تقييم مهارة المحادثة؛ حيث اعتمد المقيمون على إستراتيجيات مختلفة في تقييمهم لهذه المهارة؛ وتنقسم هذه الاختبارات إلى ثلاثة أنواع بالنظر إلى عنصر المقيّم، وعنصر المقيّم، وعنصر مقياس التقييم (Youn, 2021, pp. 107-8).

المقيّم:

وجد الباحثون أن الممارسات الشائعة لدى مقيّم مهارة المحادثة هي إما أن يقيم المتحدث مقيّم واحد، أو مقيّمان، وتجدر الإشارة إلى التوسع في انتشار الاعتماد على مقيّمين اثنين على حساب المقيّم الواحد في الآونة الأخيرة؛ لأسباب كثيرة، ربما من أبرزها الدقة، والمصادقية؛ حيث أشارت الدراسات إلى أن الاعتماد على مقيّمين يقود إلى دقة أكبر في نتائج التقييم، وبالتالي مصادقية أعلى، ويمكن أن يكون المقيمان مدرسين، أو زميلين للمتعلم (Carrió-Pastor, 2016, p. 61). وحتى يتمكن المقيمون من تقييم التحدث والتفاعل لدى متعلمي اللغة الثانية

على النحو المنشود، يجب عليهم تطوير مهاراتهم في التقييم مما يمكنهم من إصدار أحكام مهنية حول إتقان المتعلم للغة الثانية، وللمهارات التفاعلية، بما يتماشى مع معايير واضحة. وقد أظهرت الدراسات نتائج إيجابية لمستوى عال من الدقة والمصادقية في توافق تقديرات المقيمين المختلفين الذين خضعوا للتدريب، في مقابل تباين ملحوظ بين تقديرات المقيمين المبتدئين الذين لم يخضعوا للتدريب (Sandlund & Sundqvist, 2021, p. 134). إضافة إلى ما سبق، ومن أجل التوصل إلى توافق أكبر في تقديرات المقيمين، اقترح عدد من الباحثين عقد جلسات توافق بين المقيمين، تهدف إلى تضيق الفجوة بين تقديراتهم لأداء المتحدثين باللغة الثانية، حيث وجدوا أن هذه الجلسات أفادت في فهم الآليات التي يتبعها المقيمون لمنح الدرجات، وفي تطوير تعليمات تساعد على التقييم (Sandlund & Sundqvist, 2021, p. 136).

وأكدت الدراسات أيضا وجود تباين في الدرجات التي يمنحها المقيمون للطلبة في مهارة المحادثة حتى بين المقيمين الخبراء والمدرّبين تدريباً جيداً (Youn, 2021, p. 108; Al shatter, 2017)، وربما يعود هذا التباين إلى عدة أسباب، من أبرزها: التفاوت في فهم المقيمين للمحتوى، والأهواء الشخصية، والخلفية الثقافية أو الاجتماعية للمقيم. وقد حدد الباحثون عدة أبعاد قد يكون لها دور فاعل في التباين بين تقديرات المقيمين، من أبرزها: اتساق المقيّم (تبنى وتيرة ثابتة في التقييم)، وتساهله أو شدته، واستخدامه لمقياس التصنيف، وتحيزه أو تفاعله (Sandlund & Sundqvist, 2021, p. 132).

المقيّم:

وهنا انقسم التقييم إلى نوعين أيضا؛ الأول: تقييم طالب واحد فيما يعرف باختبارات التحدث الأحادية (monologic speaking tests)، والثاني تقييم طالين معا فيما يعرف باختبارات التحدث المقترن (paired speaking assessment)، أو

التفاعلي (interactive speaking assessment) (Youn, 2021, p. 107). وأشارت الدراسات إلى أن من أبرز التحديات التي يواجهها المقيّمون عند إجراء اختبارات المحادثة هي تحديات مرتبطة بالشخصية، مثل: القلق الذي يظهر على المقيّمين أثناء أداء الاختبارات، أو الانطواء؛ مما قد يؤثر على أدائهم، مما دفع تلك الدراسات إلى الحث على اتخاذ المقيّمين إجراءات تهدف إلى التقليل من مستوى القلق والتوتر الذي يشعر به المقيّمون، ومعالجة المشاكل الشخصية، مثل: الانطواء، قبل البدء بالاختبار، مثل: تقديم معلومات حول مواضيع المحادثة المستهدفة في الاختبار (Emirza & Sahril, 2021, p. 10; Köroğlu, 2019, pp. 24-25).

مقياس التقييم (Rating Scales)

يعرّف مقياس التقييم على أنه تصنيف لعدد من المستويات المستهدفة لمهارة المحادثة، يقاس على أساسه مستوى كفاية متعلمي اللغة في تلك المهارة، يقدم مواصفات إجرائية لمستويات الكفاية اللغوية. وعادة ما يتراوح في درجات التقدير بين الصفر (انعدام الكفاية) إلى درجة معينة تمثل أعلى درجات الكفاية المستهدفة. تحدد المستويات المختلفة في مقاييس التقييم بحسب قدرة المتعلم على القيام بمهام ووظائف لغوية، وإتقانه لسمات لغوية، مثل: المفردات، والنحو، والطلاقة، وترابط الأفكار، وغيرها (Fulcher, 2014, p. 88).

تستخدم مقاييس التقييم عادة لأمرين رئيسين: (١) تقييم مهارة المحادثة، و(٢) إرشاد مصممي اختبارات المحادثة لاختيار مهام معينة للاختبار. إضافة إلى استخدامات أخرى تعتمد على طبيعة المقياس، من أبرزها: (١) مقاييس موجهة للمستخدم: وتستخدم لتقديم تقرير حول الأداء النموذجي للمقيّمين، أو الأداء المحتمل في مستوى ما، (٢) مقاييس موجهة للمقيّم: تصمم لتوجيه عملية التقييم، مع التركيز على نوعية الأداء المتوقع، (٣) مقاييس موجهة للمصمم: تنتج

لمساعدة مصمم الاختبار على اختيار المهام التي سيحتوي عليها الاختبار (Fulcher, 2014, p. 89).

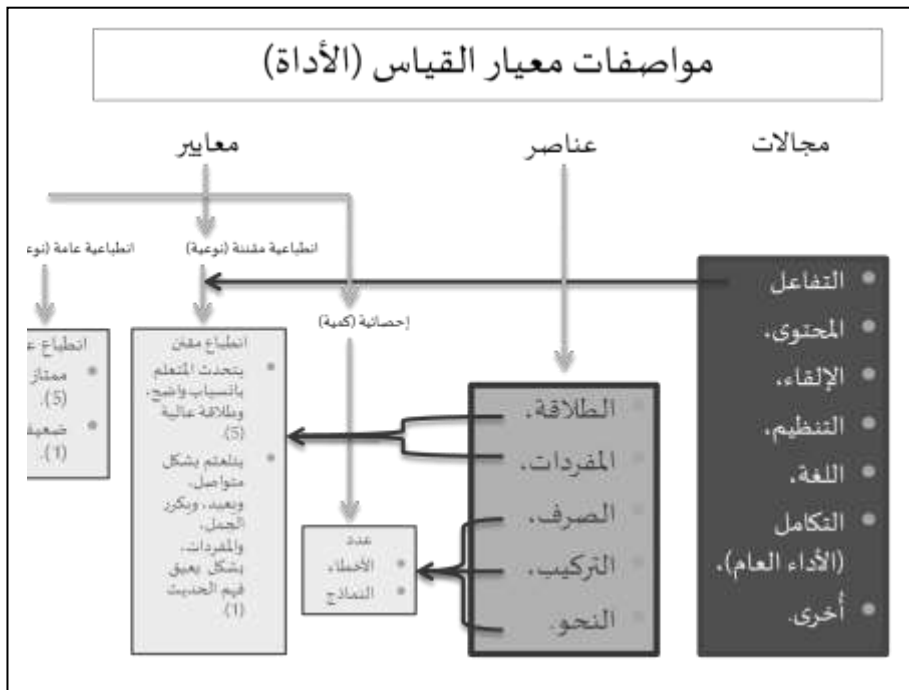
أشار الباحثون إلى أن المقيمين اعتمدوا عدة إستراتيجيات فيما يخص المقياس المستخدم في إعطاء الدرجات للطلبة المقيمين؛ بعضهم اعتمد على انطباعات شخصية مستندا إلى خبرته الشخصية في تدريس مهارة المحادثة، وآخرون اعتمدوا معيارا صمموه مسبقا لتحديد درجات المقيمين، وآخرون اعتمدوا مقاييس صممتها مؤسسات تهتم بتدريس مهارة المحادثة واختبارها، ومن أبرزها: (١) مقياس ماساشوستس لتقييم التحدث (Massachusetts Speaking Assessment Criteria)، و (٢) فيرفكس كاونتي، مدارس فرجينيا الحكومية (Fairfax County, Virginia Public schools)، و (٣) مقياس أوجدن للتحدث (Ogden Speaking Rubric)، و (٤) مقياس إلينويز للتحدث (Illinois Township District 214 Rubric). (Al shatter, 2017, p. 33)

مواصفات مقياس التقييم

مما سبق يتضح أن أية محاولة لتطوير مقياس لتقييم مهارة المحادثة عليها أن تراعي مواصفات خاصة في تصميمه، من أبرزها: (١) حصر وتحديد مجالات مهارات المحادثة التي يشملها المقياس، مثل: المحتوى، الإلقاء، التنظيم، السلامة اللغوية، الأداء العام، النبر والتنغيم، النطق السليم، المفردات، الحوار والتفاوض، وغيرها (Arung, 2016, pp. 71-72)، (٢) حصر وتحديد العناصر التي يشملها كل مجال، فعلى سبيل المثال، وفي مجال اللغة، يمكن أن يعتمد مصمم المقياس عناصر، مثل: الطلاقة، المفردات، الصرف، بنية الجملة، النحو، وغيرها (معمري، ٢٠١٨، الصفحات ٥٩٩-٦٠٠). (٣) حصر وتحديد المعايير التي يقيم على أساسها المقيّم المتحدثين؛ التي يمكن أن تقسم إلى قسمين: نوعية (انطباعية)، وكمية (إحصائية) (الشاطر، ٢٠١٩، صفحة ٧٠).

صنف بعض الباحثين مهارات التواصل الشفوي إلى عدة مجالات، من أبرزها: (١) مجال الألفاظ والتراكيب، (٢) مجال المضمون، (٣) مجال الأصوات، (٤) مجال القواعد اللغوية، (٥) مجال شخصية المتحدث، ويتضمن كل مجال من هذه المجالات عددا من الموصفات المرتبطة بمهارة التحدث، مثل: اختيار الألفاظ وتنوعها، واختيار الأفكار الملائمة للموضوع، وإخراج الحروف من مخارجها الصحيحة، وخلو الحديث من الأخطاء اللغوية، والجرأة والثقة بالنفس (الحارثي، ٢٠٢٠، الصفحات ١١٧٤-١١٧٥).

هذا، ويوضح الشكل الآتي علاقة تلك الموصفات بعضها ببعض، مع عدد محدود من الأمثلة توضح المقصود.



شكل ١: موصفات معيار القياس (الأداة)

يظهر الشكل (١) أعلاه طبيعة الترابط بين ثلاثة مكونات رئيسة في تصميم الاختبارات الشفوية، هي: المجال، والعنصر، والمعيار؛ ويقصد بالمجال هنا الإطار العام الذي يضم عددا من العناصر، مثل: المحتوى، الذي يمكن أن يضم عددا من العناصر، مثل: الترابط، والأفكار الرئيسة والفرعية، وغيرها، أو مجال اللغة، الذي يضم عناصر، مثل: الطلاقة، المفردات، الصرف، وغيرها، أما العنصر، فهو: فرع يشكل مكونا من مكونات المجال، مثل: التنغيم، النبر، الوقف، وغيرها، التي تشكل مكونات لمجال الإلقاء، والمجال الوظيفي، مثل: التعداد، الوصف، المقارنة، النقد، وغيرها. أما المعيار، فقد ذكرت الدراسة الحالية تعريفه، ومواصفاته في قسم سابق.

إلا أنه من المفيد الإشارة هنا إلى أن العناصر يمكن أن تصنف إلى قسمين: (١) عناصر يمكن قياسها كميا بالاعتماد على بيانات إحصائية، مثل: التكرار، و (٢) عناصر تقاس بناء على انطباع المقيّم، مثل: ترابط الأفكار؛ وعليه فإن المعايير تقسم بدورها إلى ثلاثة أقسام: (١) معايير كمية تعتمد على البيانات الإحصائية، مثل: قياس الطلاقة اللغوية، وهنا حددت الدراسات التي أجريت على الطلاقة اللغوية عند متعلمي اللغة الثانية عددا من المعايير المرتبطة بالطلاقة وقياسها، من أبرزها: الجوانب الزمنية لإنتاج الكلام مثل معدل الكلام، والإصلاحات، وكمية التردد وتواتره، وموقع التوقيفات، وطول فترات الكلام بطلاقة بين فترات التوقف المرتبطة بالتحليل النفسي اللغوي لجوانب الأداء والإنتاج (Mirzaei & Heidari, 2012, p. 38)، وهذه كلها يمكن برمجة الحاسب لقياسها آليا اعتمادا على الكم، و (٢) معايير انطباعية مقننة، وفيها يضع المقيّم وصفا لكل درجة يعطيها للمقيّم، ويقاس أداء المقيّم مستندا إلى نسبة تحقق الوصف في أداء المقيّم (أنظر جدول (١) أدناه لمزيد من المعلومات)، و (٣) معايير انطباعية عامة، يعتمد فيها المقيّم في منح الدرجة على رأيه الشخصي، دون تحديد مواصفات، أو سلم أداء.

الجدول (١): المعايير الانطباعية (النوعية) المقتنة

المحتوى		
الدرجة	معايير منح الدرجة	درجة المتعلم
٥	- يرتبط المحتوى ارتباطا وثيقا بمهمة المحادثة المطلوبة. - يقدم المتكلم أنواعا متعددة من المحتويات التي تلائم المهمة المطلوبة، مثل: التعميم، والتفصيل، وضرب الأمثلة، وبراهين متعددة الأشكال.	
٤	- يرتبط المحتوى ارتباطا كافيا بمهمة المحادثة المطلوبة. - يركز المتكلم على المحتوى المرتبط ارتباطا وثيقا بمهمة المحادثة. - يلتزم المتكلم بالموضوع.	✓
٣	- بعض محتوى الحديث لا يرتبط بالمهمة. - يدور المتكلم حول الموضوع.	
٢	- كثير من محتوى الحديث لا يرتبط بالمهمة. - يستخدم المتحدث مفردات ومفاهيم غير مناسبة (مثل استخدام العامية، أو مفاهيم غير ملائمة ثقافيا).	
١	- محتوى حديث المتكلم لا علاقة له بالمهمة المطلوبة. - لم يقل المتكلم شيئا.	

يظهر الجدول (١) أعلاه أن المجال المقاس هنا هو المحتوى، وفيه يشتمل العمود الأول على الدرجات التي يحصل عليها الطالب بناء على تمكنه من تطبيق العناصر المذكورة في العمود الثاني، الذي يتضمن وصفا لدرجة إجابة الطالب لكل عنصر من العناصر، في كل درجة من الدرجات في العمود الأول، ويضع المختبر إشارة في العمود الثالث مقابل الدرجة التي يستحقها المختبر. نلاحظ في الجدول (١) أن المختبر مقيد بالمواصفات المذكورة في العمود

الثاني مقابل كل درجة، ومنحه للدرجة مرهون بأداء المختبر الموضح في عامود معايير منح الدرجة (العمود الثاني). يفترض في هذا النوع من المعايير الانطباعية المقننة أن يساهم في زيادة مصداقية الاختبار وثباته؛ لأن المقيمين جميعهم ملزمون بالتقيد بمعايير منح الدرجة المنصوص عليها في العمود الثاني من الجدول. وعند قراءة الجدول (١) نعرف أن الدرجة التي منحها المقيم للمقيم هي ٤ من ٥؛ وذلك لأنه أشر مقابل الدرجة في العمود الثالث.

فوائد اختبارات المحادثة

يحتاج المعلمون إلى تحفيز الطلاب وتشجيعهم، وتزويدهم بالتدريب النفسي والطمأنينة والمشورة. ويحتاجون أيضاً إلى إيجاد طرق ودية لتصحيح أخطاء الطلاب، بحيث يصبح توفير بيئة ودية ومواتية في الفصل الدراسي أولوية (Nazara, 2011, p. 41). أظهرت الدراسات أن الطلاب يشعرون بالخجل أو قلة الحماس للتحدث في غرفة الصف، سواء في أثناء أداء مهام محادثة صفية، أو في أثناء الجلوس إلى اختبارات المحادثة (Bashir, Azeem, & Ashiq, 2011, p. 46). وعليه أصبح من الضروري لأي نظام تعليمي أن يراعي عند تعليم لغة ثانية تشخيص أداء متعلميها، من أجل التنبؤ بإمكاناتهم من جهة، ومراقبة أداء تلك النظم التعليمية، وجودة منتجاتها من جهة أخرى (Magal-Royo, Laborda, 2021, p. 251).

ولا شك في أن دعم قدرة الطلاب على التحدث بفعالية في غرفة الصف عبر الإدارة الفعالة للصف الدراسي وتعليم التحدث، يعد من المهارات الضرورية التي يجب على مدرسي اللغات الثانية إتقانها، وتطبيقها؛ حتى يصل المتعلمون إلى مستويات معقولة في التواصل باستخدام تلك اللغة الثانية، تمكنهم من التفاعل والاندماج في مجتمع تلك اللغة ببسر وسهولة (Miranda & Wahyudin, 2023, p. 41). ولا يمكن للمعلم التعرف إلى هذه القدرة دون إجراء تقييم للطلبة.

إلا أن التركيز على الاختبارات التقليدية في قياس قدرات متعلمي اللغة الثانية في مهارات القراءة، والكتابة، والفهم والاستيعاب، ليست كافية لقياس قدرات المتعلمين اللغوية؛ ذلك لأن القدرة اللغوية تعتمد على عنصرين رئيسيين: الكفاءة، والأداء. فيما تركز الاختبارات التقليدية على العنصر الأول (الكفاءة)، فإن معرفة قدرات المتعلمين على التفاعل مع الآخرين، والتواصل عن طريق اللغة الثانية، تتطلب معرفة مستوى أدائهم، وهذا لا يتحقق إلا بقياس مستوى أدائهم الشفوي (Magal-Royo, Laborda, Marçal, & Salvador, 2021, p. 252).

أكدت الدراسات ذات الصلة حول استخدام متعلمي اللغة الثانية لاستراتيجيات التعلم، أنه ولكي ينجح هؤلاء المتعلمين في تطوير قدراتهم في مهارة المحادثة، فإن عليهم أن يشاركوا بفاعلية وانتباه في عملية التعلم، وأن يستخدموا مجموعة من استراتيجيات التعلم التعويضية والإبداعية، والتقنيات التفاعلية. يشمل التحدث باللغة الثانية بطلاقة، تمكن متعلم اللغة الثانية من تطبيق قدرته اللغوية بكفاءة بشكل طبيعي مستخدماً عمليات المعالجة المعرفية الأساسية لإنتاج الكلام (Mirzaei & Heidari, 2012, p. 37)، وهنا تبرز الحاجة إلى اختبار قدرات المتعلمين في مهارة المحادثة، خاصة وأن عددا كبيرا منهم يفضلون التحدث بلغتهم الأم رغم قضائهم سنوات عديدة في تعلم لغة ثانية (Zakaria, Hashim, & Yunus, 2019, p. 3083)، فيكون الاختبار بمثابة دافع لهم لممارسة اللغة الثانية، وإيجاد الطرق والوسائل التي تمكنهم من ذلك.

هذه الجهود التي يبذلها المتعلمون من أجل ممارسة التحدث بلغة ثانية يمكن أن تعتبر جزءا مما تحدث عنه باندورا (Bandura)، في نظريته الثنائية؛ حيث أشارت النظرية إلى أن أداء الطلاب في لغة ما وإتقانهم لها يمكن أن يرتبط بالعوامل الشخصية للتحفيز التي تصنف تحت الاستراتيجية العاطفية، والعوامل البيئية التي تندرج تحت الاستراتيجية الاجتماعية (Zakaria, Hashim, & Yunus, 2019, p. 3083).

(2019, p. 3085). وتعد الاختبارات حافزا شخصيا واجتماعيا في الوقت نفسه. إضافة لما سبق فإن للاختبارات أهمية كبيرة في وضع الخطط العلاجية للتعليم، ووصف حالة الطالب في مرحلة معينة من دراسته، والكشف عن قدراته، وتمييزها، وتحديد مدى نجاح التعليم (معمرى، ٢٠١٨، صفحة ٥٩٥)، ولا شك في أن عددا كبيرا من المتعلمين يحرصون على الحصول على شهادات أو تقارير تبين مستوياتهم في المهارات اللغوية، ومنها مهارة المحادثة، ولا يتأتى ذلك إلا بإجراء اختبارات للمحادثة تمكن المعلم من تحديد مستوى المتعلم ومنحه درجة ما.

مواصفات اختبارات المحادثة:

بحسب صنوبر (٢٠١٧، صفحة ٩-١٠)، فإن اختبارات المحادثة مرت بثلاث مراحل؛ الأولى: مرحلة عشوائية تخلو من الصدق والثبات، وتعتمد على حفظ المتعلم، ولا يحدد فيها هدف للاختبار، والثانية: مرحلة علمية تستخدم فيها أساليب علمية لقياس قدرات المتعلمين، يحدد فيها الهدف، والوقت، والمكان، والثالثة: المرحلة الحديثة التي تركز على التواصل ويكون الاختبار فيها شموليا. ذكر الباحثون مواصفات عامة للاختبارات، من أبرزها: الصدق، والثبات، والموضوعية، والعملية، والحياد (ألا يميز بين المختبرين) (معمرى، ٢٠١٨، صفحة ٥٩٧)، وهي مواصفات يجب أن يتسم بها أي اختبار بغض النظر عن المهارة اللغوية التي يقيس، سواء المحادثة، أو الكتابة، أو الاستماع، أو القراءة. وتعتبر مواصفات اختبارات اللغة بشكل عام موجّهات لاختبارات المحادثة، التي تحتوي على بيان حول بنية الاختبار، ووصف للمهام التي يتكون منها الاختبار، وبيان حول نوع الإجابات التي نتوقع أن يقدمها المتقدمون للاختبار، وشرح حول كيفية تقدير الأداء. علاوة على مواصفات خاصة تتعلق بكل اختبار على حدة يجب أخذها بعين الاعتبار، ترتبط ارتباطا وثيقا بطبيعة ذلك الاختبار. ينظر إلى

مواصفات الاختبار على أنها وثائق حيوية (ديناميكية) ومتطورة يجب أن تكون مرتبطة بعملية تصميم الاختبار وتجريبه ومراجعته (Fulcher, 2014, p. 116). هذا، وعند البدء في عملية تصميم اختبارات المحادثة، فإنه من الواجب علينا مراعاة عناصر ضرورية في نجاح العملية، تتعلق بفريق التصميم، والمتقدمين للاختبار، والغاية من الاختبار، وبنية الاختبار، والمهام التي يتضمنها الاختبار، وتجريب الاختبار على عينات صغيرة، والعمل على تطوير مقاييس التقييم في مرحلة التجريب، ووضع تعليمات إدارة الاختبار، ودعم قرارات التصميم بنتائج أبحاث في المجال، وتوضيح أية قيود في تصميم الاختبار (Fulcher, 2014, pp. 117-19).

وهنا تظهر ضرورة تحقيق انسجام بين فريق التصميم من حيث تمكنهم من مجموعة من المهارات التي يتطلبها تصميم الاختبار، وامتلاكهم للمعرفة الجماعية الكافية لتصميمه، وضرورة وضع توصيف دقيق لمن يمكنه التقدم للاختبار، مرتبط بالعمر، والجنس، والخلفية الثقافية، والموقع الجغرافي، والنظام التربوي. إضافة إلى ما سبق، فإن على مصممي الاختبارات مراعاة عدد من النقاط عند تصميم اختبارات المحادثة، من أبرزها (Fulcher, 2014, pp. 120-21):

- كتابة العديد من المهام المتشابهة والمتقاربة، في المستوى، والشكل، التي يمكن وضعها في بنك المهام لإنشاء عدد من النماذج المتناظرة لاختبار التحدث.
- التحقق مما إذا كانت مهام التحدث تستهدف العمليات، أو اللغة التي أَرادها كُتاب المهام.
- تنويع شكل العناصر في الإصدارات المستقبلية للاختبار؛ بحيث يتطور الاختبار بما يتوافق مع دراسات التقييم المستقبلية، والاكتشافات الجديدة في اكتساب اللغة، واللسانيات التطبيقية.
- وفي حالة تصميم الاختبارات المحوسبة، فإن على المصمم دراسة المتاح من

التقنيات، والبرامج، التي تمكنه من تصميم اختبار ملائم لقياس مكونات الاختبار التفاعلية، والتواصلية، واللغوية، والشخصية، إضافة إلى سهولة التعامل مع الاختبار، ورضى المستخدمين عنه، وتوافر الخبرة اللازمة لديهم لاستخدامه (Magal-Royo, Laborda, Marçal , & Salvador , 2021, pp. 252-253).

وقد لاحظ الباحثون والعاملون في مجال تعليم اللغات الثانية تطوراً كبيراً في استخدام التقنيات المعاصرة في تعليم مهارة المحادثة، من ذلك: استخدام الشبكة، وبرامج التواصل الاجتماعي، مثل: الفيس بوك، واليوتيوب، وغرف الدردشة، وبعض التطبيقات الحاسوبية، مثل: web 2.0، وغيرها من التقنيات المعاصرة، التي أثرت إيجابياً على مستويات الطلبة في المحادثة، وتطوير مهارات المحادثة المختلفة لديهم (Anwar, 2021, p. 577).

مما سبق يمكننا القول بأن اختبار مهارة المحادثة، أمر ضروري لتعلمها، وممارستها بشكل سليم من قبل المتعلمين، وتظهر الحاجة الماسة إلى تصميم اختبارات آلية محوسبة خاصة بمهارة المحادثة لدى متعلميها من غير الناطقين بها؛ وذلك من أجل مواكبة التطور التقني المعاصر من جهة، ومن أجل وضع حلول للتحديات التي تواجه المقيمين عند تصميم اختبارات المحادثة من جهة أخرى.

أبرز التحديات عند اختبار المحادثة:

على الرغم من تضارب الآراء حول حدة التحديات التي تظهر عند حوسبة اختبارات المحادثة؛ حيث يرى بعض الباحثين أن هذه التحديات تمنع الاستفادة من التقنيات الحديثة في الاختبارات اللغوية، في حين يرى آخرون أن التعامل مع هذه التحديات ممكن، وأن نسبة إعاقتها للاستفادة من التقنيات في الاختبارات اللغوية يمكن السيطرة عليها، إلا أن هناك شبه إجماع على أن استخدام التكنولوجيا في الاختبارات اللغوية أصبح أمراً حتمياً لا مهرب منه، وأن المستقبل

القريب سيشهد توسعا كبيرا في هذا الاتجاه (Suvorov & Hegelheimer, 2014, p. 14). ولا تقتصر التحديات على التقنيات، وإنما هناك تحديات متعلقة بتصميم أدوات القياس، وأخرى متعلقة بمتطلبات معنوية، مثل: الوقت، ومتطلبات مادية مثل المعدات والمرافق، ومتطلبات بشرية، مثل إدارة الاختبار، وأخرى متعلقة بالاختبار نفسه، مثل: المصادقية والثبات (Aleksandrzak, 2011, p. 46).

من أبرز التحديات التقنية التي تواجه حوسبة اختبارات اللغة هو الجهد الكبير والوقت الطويل الذي تحتاجه هذه الاختبارات من أجل تأمين عدد كبير من الأسئلة لضمان اختلاف الأسئلة التي يجيب عليها المقيّمون، وعدم تكررها بشكل يؤدي إلى تسربها نتيجة حفظها، الأمر الذي يولد تحديا آخر، وهو التكلفة العالية ماديا التي يحتاجها هذا النوع من الاختبارات (Suvorov & Hegelheimer, 2014, p. 14).

إضافة إلى ما سبق يظهر تحد تقني آخر يتعلق بالاختبارات الآلية نفسها؛ حيث تساءل عدد من الباحثين عن قدرة الآلة في تقدير مستوى المقيّمين في مهارة المحادثة في مهارات متعلقة بشخصية المقيّم، مثل: قوة منطقته، وقدرته على الإقناع، وقلق التحدث، وغيرها؛ مما يجعلنا ندرك بأن اختبار مهارات المحادثة جميعها آليا أمر ليس ممكنا حتى الآن. وفي أفضل الحالات يمكننا في الوقت الحاضر تصميم اختبارات محادثة يشترك فيها العنصر البشري مع العنصر الآلي (Suvorov & Hegelheimer, 2014, p. 16).

وهناك تحد آخر متعلق بمشاكل متعلقة بقدرة المقيّمين، وهذا التحدي ينقسم إلى قسمين: مشاكل لغوية، مثل: النحو، والمفردات، والنطق، وأخرى نفسية عاطفية، مثل: القلق، والانطواء، والتوتر، والثقة بالنفس، واحترام الذات، والتعاطف، والتحفيز، والخجل، والخوف، وقلة الرغبة، التي تعد من أكثر المعوقات التي تواجه المتعلمين عند التحدث بلغة ثانية، وتستوجب حولا مسبقة، سواء أثناء

التدريس أو قبل عقد الاختبارات، وهو أمر ينطبق على اختبارات المحادثة عامة، ما كان تقليدياً، أو محوسباً (Emirza & Sahril, 2021, p. 12; Anwar, 2021, p. 572).

علاوة على ما ذكر سابقاً فإن تحدياً واقعياً يواجه تصميم وتطبيق الاختبارات المحوسبة يتمثل في التكلفة المادية العالية التي يتطلبها تصميم هذه الاختبارات، التي تحتاج إلى ميزانيات يمكن للشركات الكبرى فقط توفيرها؛ مما يجعل من الصعب على الفرق البحثية العلمية إنتاج هذه الاختبارات دون الحصول على دعم مادي، يساهم في تغطية نفقات الإنتاج التي تتطلب وقتاً، وجهداً (خضراً، وآخرون، ٢٠١٩، صفحة ٣٠).

اختبارات المحادثة المحوسبة:

هناك نوعان من التدخل الآلي في اختبارات المحادثة المحوسبة: الأول يعنى باستخدام الحاسب وسيطاً من أجل إجراء الاختبار، دون تدخل الحاسب في تقييم الطلبة، أو منحهم الدرجات؛ كأن يستخدم برنامج سكايب، أو زووم، أو غيرها، وفي هذه الحالة يكون دور الحاسب مجرد وسيط فقط، تعقد عبره المقابلة بين المقيّم والمقيّم، أما منح الدرجات والتقدير فيقوم بها المقيّم (Iwashita, Lyn, & Moore, 2021, p. 291).

أما النوع الثاني فلا يقتصر على أن يكون منصة للاختبار، وإنما يتعدى ذلك إلى تقييم المتحدثين ومنحهم درجات، وقد انتشر مؤخراً ضمن ما يسمى "Computer-assisted language testing (CALT)"، أي اختبار اللغة بمساعدة الحاسب، الذي يسعى بشكل رئيس إلى ضمان جودة الاختبار، وتحقيق العدالة بين المقيّمين، والابتكار في تصميم الاختبار (Suvorov & Hegelheimer, 2014, p. 1)، ومنه ظهرت آليات تقنية فرعية تعنى باختبار المحادثة، مثل: برامج التعرف الآلي على الحديث ("ASR" Automatic Speech Recognition)

(Doremalen, Cucchiari, & Strik, 2013)، والأنظمة الآلية لتقييم الكلام (Automated Speech Evaluation (ASE) (Suvorov & Hegelheimer, 2014, p. 1; Alsulaiman, et al., 2012).

ويلاحظ أن النوع الأول من التدخل الآلي في اختبارات المحادثة المحوسبة سهل التطبيق، ولا يحتاج إلى خبرات، أو إجراءات معقدة، فالأمر فيه يقتصر على توفير منصة سحابية عبر الشبكة لإجراء مقابلة الاختبار، في مقابل متطلبات ذات طابع احترافي معقد تلزم في تصميم النوع الثاني من التدخل الآلي الذي يُمنَح فيه المتحدثون الدرجات آلياً. هذه المتطلبات تتضمن احتياجات برمجية، ولغوية، وتربوية، وقانونية في الوقت نفسه (الشاطر، ٢٠١٩، صفحة ٧١).

وقد تزايد الاهتمام في الآونة الأخيرة بتطوير أنظمة التعرف الآلي على الكلام بغرض التقييم، وازداد عدد الدراسات المهمة بهذا الجانب الضروري في مجال تعليم اللغات، وشملت التقييم الآلي للمكتوب، والمنطوق. إلا أن التقييم الآلي للمنطوق مازال محدوداً، ويواجه مشاكل عديدة، من أبرزها: أن البرامج التي طورت لتقييم المنطوق لم تتمكن حتى الآن من التعامل مع اختلاف لكانات متعلمي لغة ما، خاصة وأن تقييم تلك البرامج للمنطوق بني على قدرات المتحدث الأصلي للغة (Suvorov & Hegelheimer, 2014, p. 13).

ولم يقتصر التقييم الآلي للمنطوق على جوانب لغوية بحتة، مثل: النطق، وسلامة اللغة، والطلاقة، وغيرها، بل تعدته إلى تقييم الكفاءة التفاعلية في الحديث (interactional competence)؛ حيث أكدت الدراسات على تزايد أهمية قياس الكفاءة التفاعلية في اختبارات المحادثة، وفيها يركز المقيّمون على مؤشرات التفاعل الحواري في حديث المختبرين، التي بدورها تزامنت مع التوسع في استخدام ما يسمى بتقنية الاختبارات المحوسبة (AI- (computer-mediated tests

(Gahtani & Al Shatter, 2012)، مثل: اختبار التوفل المحوسب، واختبار بيرسون المحوسب، التي تتضمن عناصر محادثة ضمن العناصر التي تقيسها، مما يعكس تقدماً في التقنيات المستخدمة من جهة، وواقع التواصل من جهة أخرى، في الوقت الذي أصبحت فيه التفاعلات المنطوقة من خلال أنظمة الحاسب الآلي شائعة بشكل متزايد، بالأخص وقت جائحة كورونا.

وعلى الرغم من ظهور بعض عناصر الكفاءة التفاعلية في مهام المحادثة في هذه الاختبارات (الاستجابة على سبيل المثال). إلا أن الظاهر أن تقييم أداء الطلبة يقتصر بشكل رئيس على عناصر الطلاقة، والدقة، والمحتوى، ويغفل جوانب ضرورية من عناصر الكفاءة التفاعلية، من أبرزها: تبادل الأدوار بين المتحدثين، وبدء وإنهاء الحديث (المبادرة في التحدث)، والاستجابة، والاستدراك وإصلاح الحوار (اتخاذ إجراءات لضمان استمرارية الحوار ووضوحه) (Iwashita, Lyn, & Moore, 2021, p. 285)؛ مما يعكس رؤية ضيقة للكفاءة اللغوية في مهارة المحادثة من المنظور النفسي اللغوي، ويعكس في الوقت نفسه قصوراً في شمولية تقييم الجوانب المختلفة في مهارة المحادثة لدى الطلبة (Iwashita, Lyn, & Moore, 2021, p. 283).

مما سبق يمكن القول بأن تطوير برامج تعنى بالتقييم الآلي لمهارة المحادثة، ليست بالمهمة السهلة، وتتطلب جهداً كبيراً يراعي عدداً من العناصر الضرورية، من أبرزها: مواصفات المهام المطلوبة من المتعلمين (Magal-Royo, Laborda, Marçal, & Salvador, 2021, p. 255)، ومواصفات تلك البرامج، والمعايير التي تحكم قياس المهارة، والإجراءات الأكاديمية والإدارية التي تساهم في نجاح تطوير هذا النوع من البرامج، كما يتضح أيضاً أننا لن نتمكن من تطوير برامج تقييم المحادثة آلياً، إلا إذا تمكنا من تطوير مقاييس تقييم مناسبة (Rating Scales)، وفق خطوات مدروسة من أبرزها:

١. تطوير أنظمة التعرف الآلي على الحديث (Automatic Speech Recognition-ASR)،

٢. التركيز على المعايير الكمية،

٣. العمل على إيجاد بدائل للمعايير النوعية في قياس كفاية المحادثة،

٤. تطوير إطار نظري معرفي يجمع بين اللغة والتقنية.

الخلاصة

خلصت الدراسة الحالية إلى نتيجة تشير إلى أهمية تدريس المحادثة في صفوف تعليم اللغات الثانية، وأن التدريس الناجح يركز إلى عناصر عدة من أبرزها تقييم أداء المتعلمين وقدراتهم التفاعلية والتواصلية في أثناء التحدث باللغة الثانية، وهو ما تقتصر إليه معظم فصول تدريس اللغات الثانية، بما فيها اللغة العربية. وخلصت الدراسة الحالية أيضا إلى أن تحديات تقييم مهارة المحادثة كثيرة؛ منها تدني نسبة المصادقية والثبات في الاختبارات المستخدمة من قبل المعلمين، أو المؤسسات التي تتولى تدريس اللغات الثانية.

كما خلصت الدراسة الحالية إلى أن حوسبة اختبارات المحادثة أمر ضروري وممكن، ويساهم في تجاوز عدد من التحديات التي تظهر في الاختبارات التقليدية، إلا أن تحديات جديدة متعلقة بالاختبارات المحوسبة الآلية تحتاج إلى تركيز، ومزيد من الدرس، ومن أبرزها قياس قدرات المتعلمين المبنية على مجالات نوعية تعتمد على معايير انطباعية لا يمكن للآلة التعامل معها، وتحتاج إلى تدخل العامل البشري بشكل مباشر، ويصعب تحويلها إلى معايير كمية.

ووضعت الدراسة الحالية تصورا يتضمن المجالات والمعايير التي تبنى على أساسها أدوات قياس قدرات المقيمين في مهارة المحادثة، التي تشكل خطوة ضرورية قبل البدء في تصميم اختباراتهما. وتجدر الإشارة هنا إلى عدد من الخطوات التي يمكن اتباعها عند تقويم مهارة المحادثة، من أبرزها:

- وضع إطار نظري مرجعي (لغوي، تربوي، تقني)،
 - تطوير معايير القياس (لغوية، تربوية، تقنية)،
- وفي الاختبارات الآلية:
- التوفيق بين قدرات البرامج الحاسوبية، وبين معايير التقييم الشفوي اللغوية،
 - تصميم البرامج بالاشتراك بين المبرمجين واللغويين،
 - تجربة البرامج قبل تعميمها،
 - متابعتها بعد التعميم.
- إضافة إلى ما سبق، خلصت الدراسة إلى ضرورة العمل على تحويل المعايير الانطباعية بنوعها المقنن وغير المقنن إلى معايير كمية؛ حتى يتسنى لمصممي الاختبارات برمجتها آلياً. الأمر الذي يحتاج إلى تضافر جهود اللغويين والتربويين والحاسوبيين معاً.

قائمة المراجع العربية والأجنبية

- جيدور، ع. (2021). تدريس مهارة المحادثة من النظرية إلى التطبيق: ملاحظات بخصوص المعرفة البيداغوجية القاعدية لأساتذة اللغة العربية، *مجلة الباحث*، 173-149، (1) 13،
- الحارثي، ع. م. (2020). تقويم مستوى الأداء الشفهي لدى طالبات المرحلة الثانوية في ضوء مهارات التحدث والوعي بإستراتيجياته، *المجلة التربوية*، 73، 1157-1241.
- خضر، م. ز، السعودي، م.، مجد، صوالحة، م.، عابنة، س، حمدان، ي، الخطاب، م. (2019) *دليل أبحاث حوسبة اللغة العربية*، الجزء الأول، عمان، الأردن: مجمع اللغة العربية الأردني.
- معمري أ. (٢٠١٨). أسس وضع اختبارات تعليم اللغة العربية للناطقين بغيرها. *مجلة الباحث في العلوم الإنسانية والاجتماعية*، ٣٥، ٦٠٧-٥٩٥.
- الشاطر، غ. (٢٠١٩). اختبارات قياس الكفاءة اللغوية لغير الناطقين باللغة العربية: التطبيق والمواصفات والمقاييس. اللسانيات، ٢٥ (١).
- صنوبر، خ.، والنصراوي، م. (٢٠١٧). *تطوير اختبار الكفاية اللغوية لمتعلمي اللغة العربية الناطقين بغيرها*، رسالة ماجستير غير منشورة، جامعة عمان العربية، عمان.

المراجع الأجنبية

- Al shatter, G. (2017). Validity of Oral Tasks Testing in Arabic L2 Teaching. *Linguistica Communicatio*, 18(1-2), 27 – 43.
- Aleksandrzak, M. (2011). Problems and challenges in teaching and learning speaking at advanced level. *Glottodidactica*, 37, 37-48. Retrieved 2024, from <https://doi.org/10.14746/gl.2011.37.3>
- Al-Gahtani, S., & Al Shatter, G. (2012). Multimethod Approach to Research in Interlanguage Pragmatics. *World Academy of Science, Engineering and Technology* (71).
- Alsulaiman, M., Ghulam , B. M., Al-Gabri, M., Ali, Z., Al Shatter, G. H., & Al-Kahtani, S. A. (2012). Identification of Arabic L2 Learners Origin. *International Symposium on Automatic Detection of Errors in Pronunciation Training* (pp. 107-113). Stockholm, Sweden: KTH.
- Altun, M., & Sabah, R. (2020). The Effect of Cooperative Learning Strategies in the Enhancement of EFL Learners' Speaking Skills. *Asian EFL Journal*, 27(2), 144-171.
- Anwar, M. S. (2021). The Most and the Least Enhanced Speaking Skill (Fluency, Accuracy of Grammar, and Accuracy of Vocabulary) by a Web 2.0 - Enhanced Project-Based Learning Instructional Program of the Jordanian EFL Female Eleventh Grade Students. *IUGJEPS*, 29(1), 571-589.
- Arung, F. (2016). Improving the Students' Speaking Skill through Debate Technique. *Journal of English Education*, 1(1), 70-77.
- Bashir, M., Azeem, M., & Ashiq, D. (2011). Factor Effecting Students' English Speaking Skills. *British Journal of Arts and Social Sciences*, 2(1).
- Carrió-Pastor, M. L. (2016). Should peer assessment be included in

- foreign language testing? The role of motivation in testing. In M. L. Carrió-Pastor (Ed.), *New Challenges for Language Testing: Towards Mutual Recognition of Qualifications*. Newcastle Upon Tyne, UK: Cambridge Scholars Publishing.
- Doremalen, J. v., Cucchiarini, C., & Strik, H. (2013). Automatic pronunciation error detection in non-native speech: The case of vowel errors in Dutch. *Acoustical Society of America*, 134(2), 1336–1347.
- Emirza, F., & Sahril, M. (2021). Investigation Of English Speaking Skills Performance Of Introvert Students In Speaking Class. *English Journal*, 15(1), 10-17.
- Fulcher, G. (2014). *Testing Second Language Speaking*. New York, USA: Routledge Taylor and Francis Group,.
- Iwashita, N., Lyn, M., & Moore, P. (2021). Operationalising Interactional Competence in Computer-mediated Speaking Tests. In M. Salaberry, & A. R. Burch (Eds.), *Assessing Speaking in Context Expanding the Construct and its Applications*.
- Köroğlu, Z. Ç. (2019). Interventionist Dynamic Assessment's Effects on Speaking Skills Testing: Case of Elt Teacher Candidates. *Advances in Language and Literary Studies*, 10 (3), 23-31.
- Kuśnierek, A. (2015). Developing students' speaking skills through. *World Scientific News*, 1, 73-111.
- Magal-Royo, T., Laborda, J. G., Marçal , C. M., & Salvador , A. S. (2021). Alternative Computer Assisted Communicative Task-based Language Testing: New Communicational and Interactive Online Skills. *iJET*, 16(19), 251-260.
- Miranda, J. A., & Wahyudin, A. Y. (2023). Pre-Service Teachers' Strategies in Improving Students' Speaking Skills. *Journal of English Language Teaching and Learning (JELTL)*, 4(1), 40-47.
- Mirzaei, A., & Heidari, N. (2012). Exploring L2 Speech-Production

- Management from a Cognitive Perspective: A Focus on Fluency, *Iranian Journal of Applied Linguistics*. IJAL, 15(2).
- Nazara, S. (2011). Students' Perception on EFL Speaking Skill Development. *Journal of English Teaching*, 1(1).
- Pratiwi, Z. F., & Ayu, M. (2020). The Use of Describing Picture Strategy to Improve Secondary Students' Speaking Skill. *Journal of English Language Teaching and Learning (JELTL)*, 1(2), 38-43.
- Rizqiningsih, S., & Hadi, M. S. (2019). Multiple Intelligences (MI) on Developing Speaking Skills. *English Language in Focus (ELIF)*, 1(2), 127-136.
- Sandlund, E., & Sundqvist, P. (2021). Rating and Reflecting: Displaying Rater Identities in Collegial L2 English Oral Assessment. In B. M. Rafael Salaberry and Alfred Rue Burch (Ed.), *Assessing Speaking in Context Expanding the Construct and its Applications* (p. 149). Blue Ridge Summit: Multilingual Matters, [2021].
- Suvorov, R., & Hegelheimer, V. (2014). Computer-Assisted Language Testing. In A. J. Kunnan (Ed.), *The Companion to Language Assessment* (1 ed.). John Wiley & Sons, Inc.
- Torky, S. A. (2006). *The Effectiveness of a Task- Based Instruction program in Developing the English Language Speaking Skills of Secondary Stage Students*, Unpublished Master's thesis. Egypt: Ain Shams University.
- Youn, S. J. (2021). Investigating Raters' Scoring Processes and Strategies in Paired Speaking Assessment. In B. M. Rafael Salaberry and Alfred Rue Burch (Ed.), *Assessing Speaking in Context Expanding the Construct and its Applications*. Blue Ridge Summit: Multilingual Matters, [2021].
- Zakaria, N., Hashim, H., & Yunus, M. (2019). A Review of Affective Strategy and Social Strategy in Developing Students' Speaking Skills. *Creative Education*, 10, 3082-3090.