

اختبارات قياس الكفاءة اللغوية لغير الناطقين باللغة العربية: التطبيق والمواصفات والمقاييس

غسان بن حسن الشاطر
جامعة نزوى، سلطنة عمان
alshatter@gmail.com

تاريخ الاستلام: 2018/10/30 تاريخ القبول: 2019/01/28

الملخص

تسلط هذه الورقة الضوء على قضية ذات علاقة مباشرة بتعليم العربية لغير الناطقين بها، ألا وهي اختبارات قياس الكفاءة اللغوية، ويقصد بذلك تلك الاختبارات التي تعنى بتحديد كفاءة متعلم اللغة العربية من غير الناطقين بها في مهارة لغوية، مثل: الاستماع، والقراءة، والمحادثة، والكتابة، أو عنصر لغوي، مثل: النحو، والمفردات، والأصوات، والثقافة، أو متطلب لغوي، مثل: البلاغة، والأدب، والعروض، وغيرها، من أجل أداء مهمة مستقبلية ليست مرتبطة بمقرر دراسي معين.

بحسب علم الباحث يمكن حصر عدد متواضع من هذه الاختبارات على مستوى العالم العربي، أو المستوى العالمي، وسيتم التطرق إليها بإيجاز في ثانيا هذه الورقة البحثية، التي تحاول مناقشة المواصفات والمقاييس التي تبنتها الهيئات المشرفة على تطوير هذه الاختبارات، ووضع وصف لكل اختبار بما في ذلك أهدافه ومحتواه وفلسفته.

ستركز الورقة على وصف اختبار جامعة الملك سعود لتحديد الكفاءة اللغوية لدى غير الناطقين بالعربية الذي طوّر في أثناء السنوات من 2012م وحتى 2014م؛ وذلك لأن الباحث شارك في تطوير هذا الاختبار عندما كان أستاذاً لكرسي تعليم العربية لغير الناطقين بها في جامعة الملك

المؤلف المراسل: غسان بن حسن الشاطر، البريد الإلكتروني: alshatter@gmail.com

سعود. كما ستركز الورقة أيضا على عرض نتائج تجربة عملية لقياس المصدقية في اختبارات قياس الكفاءة الشفهية لدى متعلمي العربية من غير الناطقين بها.

إضافة إلى ما ذكر سابقا، تتطرق هذه الورقة إلى عرض جوانب القوة، وجوانب الضعف في هذه الاختبارات، ومناقشة أبرز الصعوبات التي قد تواجه عملية تطويرها وبعد ذلك تنفيذها. وتخلص إلى وضع تصور يحدد أبرز الحاجات اللازمة لتطوير اختبار دولي لقياس كفاءة متعلمي اللغة العربية من غير الناطقين بها في مهاراتها، وعناصرها، ومتطلباتها.

الكلمات المفتاح:

اختبارات الكفاءة - اللغة الثانية - مواصفات - مقاييس - حوسبة اللغة.

Epreuves de mesure de la compétence linguistique pour les non arabophones: Applicabilité, spécifications et normes

Résumé

Ce papier met en lumière une question en relation directe avec l'apprentissage de la langue arabe pour les non arabophones. Cette question concerne les épreuves de mesure de la compétence linguistique. Plus précisément, ces épreuves permettent d'estimer l'aptitude de l'apprenant non arabophone de la langue arabe, dans une habilité linguistique comme : l'écoute, la lecture, la discussion, l'écriture ; ou un élément linguistique comme la grammaire, le vocabulaire, les sons, la culture ; ou encore une exigence linguistique comme la littérature, la métrique poétique ou autre; dans la perspective d'une activité future non liée à un enseignement académique précis.

A la connaissance de l'auteur, un nombre modeste de ces épreuves existent au niveau du monde arabe, voire mondial, lesquelles feront l'objet de cette recherche, qui tente la discussion des spécifications et des normes adoptées par les diverses institutions supervisant le développement de ces épreuves ; et aussi de décrire chaque épreuve tout en précisant son contenu, son but et sa philosophie.

Le papier focalise sur la description de l'épreuve de l'université du Roi Al Saoud pour la détermination de la compétence linguistique en langue arabe pour le non arabophone, épreuve développée durant les années 2012-2014, en raison de la participation de l'auteur au développement de cette épreuve lorsqu'il était maître de la chaise des recherches de l'enseignement de l'Arabe pour les non arabophones, à l'université du Roi Al Saoud. Le papier focalise également sur la présentation des résultats d'une expérience pratique pour la mesure de la crédibilité dans les épreuves de l'estimation des compétences orales chez les apprenants non arabophones de l'Arabe.

En outre, nous aborderons dans ce papier les points forts et les points faibles

dans ces épreuves et la discussion des principales difficultés qui peuvent entraver leur développement ainsi que leur exécution par la suite. Nous concluons par l'élaboration d'une conception déterminant les principaux besoins pour le développement d'une épreuve internationale pour la mesure de la compétence chez les apprenants non arabophones de l'Arabe, qu'il s'agisse de ses habilités, de ses éléments ou encore de ses exigences.

Mots clés:

Épreuves de compétence - langue seconde - spécifications - normes - linguistique computationnelle.

Arabic L2 Proficiency Test: Practice, Specifications and Standards

Abstract

This paper focuses on Arabic L2 proficiency tests that measures Arabic L2 learners' proficiency in the language skills, such as: listening, reading, speaking and writing; elements, such as: syntax, vocabulary, phonetics, and culture; and language requirements, such as: genres, prosody, and rhetoric.

To the knowledge of the researcher, few tests currently exist either through the Arab world or worldwide. This paper will briefly discuss these tests, especially the specifications and standards that the developers consider to design them including the test objectives, content, and rationale.

The current paper pays extra attention to the King Saud University Arabic L2 Proficiency Test which was developed during the years 2012 through to 2014. The researcher involvement in the development of the test while working as head of TASOL Research Chair at King Saud University allows for a comprehensive understanding of the test. In addition, the paper will highlight the results of an experiment that the researcher conducted to examine the validity of Arabic L2 oral tasks tests.

The current paper also sheds light on the strengths and areas of development of Arabic L2 proficiency tests and concludes with a framework that identifies the most necessary needs to develop an international Arabic L2 proficiency test to measure Arabic L2 learners' proficiency including skills, elements and requirements.

Keywords:

Proficiency tests - second language - standards - CALL.

مقدمة

يقيس اختبار الكفاءة اللغوية قدرة المتعلم اللغوية في ضوء خبراته المتراكمة السابقة ومدى استعدادده للقيام بمهام لغوية مختلفة في المستقبل، وهو لا يعتمد في محتواه على أي مقرر أو برنامج دراسي معين (عبد الخالق، 1996، ص38)، بل يرتبط بمستوى معين يعرف بالمعيار أو الضابط المعياري (AZU, 2007)، ويقيس مهمة معينة يطلب من المتعلمين أداؤها في المستقبل، على سبيل المثال: الالتحاق ببرنامج جامعي معين، أو الالتحاق بوظيفة ما، ويقدم رؤية عامة حول معرفة المتعلمين وقدراتهم اللغوية في اللغة الإنجليزية على سبيل المثال يمكن اعتبار اختباري التوفل (TOEFL)، والآيلتس (IELTS) مثالين شهيرين على هذا النوع من الاختبارات (Ezekiel, 2016).

وقد ازدادت الحاجة في الآونة الأخيرة إلى هذا النوع من الاختبارات نظرا للانتشار السريع لتعلم العربية من قبل غير الناطقين بها في معظم أرجاء العالم، والتحاقهم ببرامج لغوية مختلفة في دول عربية وغير عربية؛ مما أدى إلى سعي عدد من الجهات المهمة بتعليم العربية لغير الناطقين بها إلى بناء اختبار كفاءة يقيس قدرات المتعلمين اللغوية تمهيدا لإلحاقهم في برامج لغوية معينة، أو من أجل التقدم لوظيفة ما في العالم العربي تتطلب معرفة باللغة العربية.

الإشكالية

ظهرت صعوبات عديدة في أثناء بناء اختبارات الكفاءة وتطبيقها، سيتعرض البحث الحالي إلى بعض منها في أثناء عرضه لوصف مقتضب لبعضها في الدول العربية وخارجها، متطرقا إلى المتطلبات الضرورية لبناء هذا الاختبار، عارضا تجربتين للباحث؛ واحدة تتعلق بمشاركته في تطوير اختبار قياس الكفاءة اللغوية لغير الناطقين بالعربية في جامعة الملك سعود، والأخرى تتعلق بدراسة أجراها الباحث حول مصداقية اختبارات المحادثة الشفهية، إضافة إلى ما سبق سيناقش البحث أبرز المواصفات والمقاييس الواجب توافرها في هذا النوع من الاختبارات.

العرض

قسم البحث هذه الاختبارات من حيث الجهة المستهدفة إلى ثلاثة أنواع:

- (1) اختبارات قياس الكفاءة لغير الناطقين بالعربية،
- (2) اختبارات قياس الكفاءة للناطقين بالعربية ولغير الناطقين بها،
- (3) اختبارات قياس الكفاءة للناطقين بالعربية، وفيما يلي عرض موجز لكل منها:

1. اختبارات الكفاءة لغير الناطقين بالعربية

1.1. العرفان - اختبار اللغة العربية للناطقين بغيرها

بحسب مصممي الاختبار، فإنه يعتمد على الإطار المرجعي الأوروبي (CEFR) في تحديد المستويات اللغوية، وعلى نظام البرومتك في تطوير الأسئلة، ويهدف إلى قياس قدرات المختبرين ممن أمضوا سنتين فأكثر في دراسة العربية وأصبحوا في المستوى المتقدم. يقيس الاختبار خمسة مجالات، هي: الكتابة، والقراءة، والاستماع، والتحدث، واللغة. ومن أهدافه أيضا إصدار شهادة معتمدة من جهة رسمية في العالم العربي وجهات معنية في العالم لتعلمي العربية من غير الناطقين بها تبين قدرتهم اللغوية استعدادا لقبولهم في عمل، أو برنامج دراسي (العرفان، 2014).

2.1. اختبار إجادة اللغة العربية - الأكاديمية العربية في الإنترنت

يقيس اختبار الأكاديمية العربية في الإنترنت خمسة مجالات، هي: فهم المسموع، والقراءة، وبنية الجملة، والكتابة، والمحادثة. طور الاختبار مجموعة من الخبراء المتخصصين سنة 2002م، وبحسب الأكاديمية العربية فإن الاختبار يعتبر الاختبار الدولي المعياري لقياس الكفاءة في اللغة العربية فيما يخص المصادقية والثبات. صمم الاختبار بشكل مرن بحيث يتكيف مع مستوى المتعلم باستخدام الحاسوب في المراحل الثلاث الأولى (الاستماع، والقواعد، والكتابة)، بينما يعتمد القسم الرابع (الكتابة) على نمط الأسئلة المقالية المفتوحة (open ended)، في حين يجري القسم الخامس (المحادثة) مقيم مختص. يقيم كل جزء عددا مختلفا من

المهارات اللغوية، ويختلف الاختبار في مدته بحسب عدد الأسئلة اللازمة لتحديد مستوى المتعلم بدقة.

وبحسب مطوري الاختبار فإن أبرز ميزاته، هي (Arab-Academy, 2015):

- النموذج الشبكي لاختبار قياس الكفاءة في اللغة العربية،
- يحدد مستوى المختبرين في اللغة العربية بدقة،
- يساهم في تعزيز فاعلية برامج تعليم العربية باستخدامه اختبارا قريبا وبعديا،
- يقيم المختبرين قبيل تسجيلهم في دورات اللغة العربية أو الجامعات العربية،
- يسهل عملية التسجيل في برامج اللغة العربية التي تتطلب حدا أدنى من الكفاءة اللغوية في اللغة العربية،
- يزود الموظفين بدليل يوضح مستوى كفاءتهم في اللغة العربية لتقديمه للجهة الموظفة في حال طلبها.

3.1. اختبار الكفاءة في اللغة العربية لغير الناطقين بها، المركز الوطني

للقياس والتقويم في التعليم العالي - المملكة العربية السعودية

يقيس الاختبار مستوى كفاءة اللغة العربية لدى غير الناطقين بها، وبحسب المطورين فإنه مصمَّم وفق المعايير والأسس العلمية المتعارف عليها عالميا في مجال القياس والتقويم، ويمكن أن يستخدم في المجالات الآتية (المركز الوطني للقياس، 2014):

• الانضمام إلى برامج اللغة أو المسارات الأكاديمية التي تتطلب مستويات لغوية معينة،

• الإعفاء من بعض المتطلبات الأكاديمية،

• الالتحاق بالوظائف التي تتطلب معرفة اللغة العربية،

• توثيق المستوى اللغوي لدارس اللغة أثناء الدراسة وبعدها.

يقيس الاختبار كفاءة المختبرين في مهارات الاستماع والقراءة والكتابة ويتكون

الاختبار من المهارات الآتية بالنسب الموضحة: فهم المقروء 40 %، والكتابة 35 %،

وفهم المسموع 25 %، كما يتوزع الاختبار على ستة أقسام، مدة كل قسم (25) دقيقة، والجدير بالذكر أن نسبة ثبات الاختبار بلغت في نسخته الأولى 94 %؛ مما يعني أن هذا الاختبار يتمتع بنسبة ثبات عالية.

4.1. اختبار الكفاءة في اللغة العربية للناطقين بغيرها - الجامعة الإلكترونية السعودية -

يهدف اختبار اللغة العربية المعياري لغير الناطقين بها في الجامعة الإلكترونية إلى قياس مهارات متعلمي اللغة العربية من غير أبنائها، في مختلف المهارات التي تقوم عليها الاختبارات المعيارية العالمية: القراءة، والاستماع، والكتابة، والتحدث، واللغة، وذلك لقياس الطلاقة اللغوية ومدى تمكن متعلمي اللغة العربية من مختلف الجوانب اللغوية خلال مرحلة دراستهم في برنامج تعليم اللغة العربية لغير الناطقين بها (الجامعة السعودية الإلكترونية، 2017).

5.1. اختبار الكفاءة في البرنامج الخاص لتعليم اللغة العربية - جامعة مولانا مالك إبراهيم الإسلامية الحكومية بملانق، أندونيسيا -

يقيس الاختبار مهارة الاستماع ومهارة القراءة والقواعد ومهارة الكتابة، ويهدف إلى منح شهادة الكفاءة في اللغة العربية للذين يهتمون بهذا الأمر. وتستخدم هذه الشهادة إما للتوثيق أو للتقديم لمنحة دراسية (جامعة مولانا مالك، 2015).

6.1. اختبار الكفاءة في اللغة العربية-جامعة ليبينج

طور الاختبار في معهد الاستشراق في جامعة ليبينج، وقيس المهارات الأربع (الاستماع، والمحادثة، والقراءة، والكتابة)، إضافة إلى قياس مهارة المتعلمين في استخدام الشبكة العنكبوتية، والرسائل النصية، والبريد الرقمي، ومواقع التواصل الاجتماعي، مثل: الفيسبوك، وتويتر، ويوتيوب، وغيرها.

يعتبر مطور الاختبار (اكهاردتشلز، Ecjehard Schulz) أن أحدا لا يتحدث العربية الفصيحة، وعليه فإن اختبار المحادثة سيقاس مهارة المحادثة في اللهجات الآتية مقسمة بحسب المناطق إلى أربع فئات:

1. الجزيرة العربية، والعراق، والخليج العربي،

2. سوريا، ولبنان، وفلسطين،

3. مصر، وليبيا، والسودان،

4. المغرب، والجزائر، وتونس.

ويقسم الاختبار إلى ثلاث مراحل بحسب المستويات اللغوية المقررة في المعيار الأوروبي (CEFR)، أما القضايا والمواضيع التي يستخدمها الاختبار في قياس مهارات المختبرين فمعظمها وظيفي ومتعلق بالحياة اليومية للمختبر (Al-Arabiyya, 2016).

7.1. اختبار جامعة نيويورك لقياس كفاءة متعلمي اللغة العربية

يتكون الاختبار من أربعة مجالات، هي: الاستماع، والتعبير، والترجمة إلى الإنجليزية، والترجمة من الإنجليزية، أما الاستماع فيتكون من نص مسموع متبوع بأسئلة يجاب عليها كتابيا، وأما التعبير فهو عبارة عن كتابة حوالي مئة وخمسين كلمة حول موضوع معين، ثم بعد ذلك يطلب من المختبر أن يترجم نصا من العربية إلى الإنجليزية، ونصا آخر من الإنجليزية إلى العربية، في مدة أقصاها ساعتين. يحصل المختبر على درجته من اثنتي عشرة نقطة، بحيث تشكل كل أربع نقاط مستوى، ويكون مجموع المستويات ثلاثة، ويمكن للمختبر أن يعيد الاختبار مرة كل مئة وثمانين يوما (Bellevue, 2016، بليفيو).

8.1. اختبار بيرلنك سوفتوير، الولايات المتحدة الأمريكية (-Pearlink Sof-tware)

يسمح هذا الاختبار للمختبرين بالتعرف إلى قدراتهم اللغوية في مجالات الفهم والاستيعاب بالاعتماد على برامج تعليمية معينة، وقد صمم على شكل ألعاب بسيطة تتضمن أسئلة اختيار من متعدد. ومن أهداف الاختبار أيضا تذكير المختبرين بما درسوه في برنامج اللغة العربية، وينصح مصممو الاختبار به قبل أداء اختبار كفاءة حقيقي، أو الالتحاق ببرنامج لغة عربية في العالم العربي.

يتكون الاختبار من أكثر من ألف سؤال مع أكثر من خمسة آلاف إجابة "صح"

أم "خطأ"، وهي تقيس مهارات اللغة العربية في المستويين المتوسط والمتقدم (Pear-link Software, 2016).

9.1 اختبار سبراتش كافي لانجوجز بلس، كندا (Sprachcaffe Languages) (Plus)

يقيس هذا الاختبار ثلاثة مجالات، هي: القراءة، والكتابة، والقواعد، ويتكون من ثمانية وخمسين سؤالاً، تتراوح إجاباتها بين الاختيار من متعدد وملء الفراغات، ليس له مدة محددة، ويجرى كاملاً باستخدام الحاسوب، ويحصل المختبر على نتيجته فور الانتهاء من الاختبار (Sprachcaffe, 2016، سبارتشكافي).

10.1 امتحان الكفاءة العالمي للغة العربية للناطقين بغيرها - جامعة النجاح الوطنية - فلسطين

بدأ تطبيق الاختبار الفعلي في سنة 2014م، وهو مصمم لاختبار متعلمي العربية من غير الناطقين بها، وقد اعتمد الاختبار من قبل البرلمان اليوناني، ويهدف إلى:

- قياس الكفاءة اللغوية في اللغة العربية بطريقة معيارية،
- مدّ جسور التواصل اللغوي والثقافي مع شعوب العالم المختلفة،
- نشر اللغة العربية في العالم،
- تشجيع الاهتمام باللغة العربية وتعلمها والإقبال عليها.
- وفيما يلي أبرز ما يتصف به هذا الامتحان:
- يعتمد الامتحان الإطار المرجعي الأوروبي المشترك للغات أساساً في اختبار المتقدمين للامتحان،
- يقسم الامتحان مستويات اللغة إلى ستة مستويات: A1، A2، B1، B2، C1، C2،
- يقسم الامتحان إلى مهارات: فهم المسموع، وفهم المكتوب، والتعبير الكتابي، والتعبير الشفهي، وتعطى كل مهارة 25% من الدرجة الكلية (جامعة النجاح، 2014).

2. اختبارات الكفاءة للناطقين بها ولغير الناطقين بها

1.2. الجامعة الأمريكية في القاهرة

يقدمه مركز التعليم المستمر في الجامعة الأمريكية في القاهرة، ويقسم إلى ثلاثة أقسام، هي: (1) اختبار الكفاءة في اللغة العربية (APT)، و(2) اختبار الكفاءة في اللغة العربية الفصحى (PTMSA)، و(3) اختبار الكفاءة في اللغة العربية العامية (PTSAC)، ويقدم للأفراد والمؤسسات (الجامعة الأمريكية بالقاهرة، 2016).

2.2. اختبار الكفاءة في اللغة العربية في الجامعة الأردنية

يستهدف اختبار الكفاءة اللغوية طلبة الدراسات العليا المقبولين في الجامعة الأردنية، وتتضمن مجالات الامتحان:

- (1) كفاية القدرة على فهم المقروء واستيعابه وتحليله ونقده،
 - (2) سلامة التعبير وفق معايير الصواب اللغوي ومراعاة إخراج الخطاب المكتوب إخراجاً متسقاً،
 - (3) أداء النص المكتوب أداءً صحيحاً معبراً.
- ويقسم الطلبة إلى ثلاث فئات، تكون علامة النجاح في هذا الامتحان لكل فئة على النحو الآتي:

- (1) الطلبة الذين يدرسون باللغة العربية ولغتهم الأم هي اللغة العربية: 70 %،
- (2) الطلبة الذين يدرسون باللغة العربية ولغتهم الأم غير العربية: 60 %،
- (3) الطلبة الذين يدرسون بلغات أخرى: 60 % (الجامعة الأردنية، 2016).

3.2. اختبار شهادة الكفاءة الدولية في اللغة العربية، (التنال العربي 2015)

اختبار موجّه إلى الناطقين باللغة العربية، وغير الناطقين بها الذين يدرسون في معاهد وجامعات وكليات تدرس باللغة العربية؛ لأغراض أكاديمية أو وظيفية، ويركز على الجانب الوظيفي للغة؛ أي اللغة المستخدمة في الحياة اليومية التي تمكّن المتعلّم من ممارستها ممارسة صحيحة، وعلى الجانب المنطوق من اللغة (مهارة الاستماع والخطاب)، وعلى الجانب اللغوي المكتوب، أي فهم اللغة حين تسمع

وحين تكتب وتقرأ، وإفهامها للآخر كلاماً وكتابة (التنال، 2015).

3. اختبارات الكفاءة للناطقين بالعربية

1.3. اختبار العين لقياس الكفاءة اللغوية (2012م)

يهدف اختبار العين إلى قياس كفاءة أبناء العربية في التواصل باللغة الفصيحة (اللغة الأم)، وقد صُممت المنظومة وفق المعايير العالمية المعتمدة في بناء اختبارات الكفاءة اللغوية. ويستهدف جميع القائمين بأعمال أو وظائف أو مهام تتطلب منهم التمكن من اللغة العربية، مثل: المدرسين (الذين يستخدمون اللغة العربية في التدريس)، والمحامين، وخطباء المساجد، والصحفيين، والإعلاميين، والقضاة، والدبلوماسيين، والمترجمين، والقائمين بأعمال (السكرتارية)، والإداريين، وغيرهم من فئات المجتمع، كما يستهدف أيضاً الطلبة والدارسين والمتقدمين إلى العمل، في بيئات تكون العربية لغة الاتصال فيها.

أما المجالات التي يقيسها الاختبار، فهي:

(1) الشفهي: استماعاً وتحديثاً،

(2) التحريري: قراءة وكتابة.

اعتمد المطورون على وثيقة كفايات الاختبار التي صممت بعد دراسة وثائق كفايات اللغة العربية في مراحل التعليم العام (قبل الجامعي) في عدد من الدول العربية، وقد استأنس الاختبار (بحسب مطوريه) بأشهر المعايير العالمية المعتمدة في تعليم اللغات، وقياس كفاءة المتواصلين بها. قسمت المستويات التي يقيسها الاختبار إلى (9) مستويات متدرجة، وقد خضع لعمليات تقييم وتحكيم من شخصيات ومؤسسات علمية مرموقة، كما تمت تجربته أكثر من مرة للتحقق من الصدق والثبات والموضوعية والواقعية (جامعة الإمارات العربية المتحدة، 2012).

4. اختبار جامعة الملك سعود لقياس الكفاءة اللغوية لغير الناطقين بالعربية

استهدف الاختبار غير الناطقين بالعربية، وبذل في تطويره جهد كبير؛ حيث كان

أحد مشاريع كرسي أبحاث تعليم العربية لغير الناطقين بها في جامعة الملك سعود، حيث بدأ العمل على تطويره سنة 2011م، وتمكنت اللجنة المسؤولة عن عملية التطوير من وضع ما يقرب من ألفي سؤال تتناول أربعة مجالات لقياس كفاءة متعلمي العربية الفصيحة من غير الناطقين بها، هي: الاستماع، والقراءة، والكتابة، والمحادثة؛ بحيث يقاس المجالان الأول والثاني إما ورقياً أو بوساطة الحاسوب، أما المجال الثالث فيستخدم فيه الحاسوب في الكتابة، إلا أن التقييم يجري من قبل مقيم، ويقاس المجال الرابع (المحادثة) كاملاً بوساطة مقيم. وقد استند الاختبار إلى معايير أكتفل (المجلس الأمريكي لتعليم اللغات الأجنبية، ACTFL) الأمريكية في تقييم اللغات، والتي تناولت اللغة العربية من بين عدد من اللغات التي تدرس في الولايات المتحدة الأمريكية لغة ثانية. ويتسم الاختبار بأنه معياري مقنن ويختبر طيف الكفاءة بأسره، من المبتدئ إلى المتوسط، فالمتقدم، وأخيراً المتميز. وقد حدد لكل سؤال أربعة بدائل متشابهة بالطول وتبدو لأول وهلة وكأنها كلها صحيحة، كما استعين بالصور المعبرة في مرحلة المبتدئ. تتدرج الأسئلة ضمن كل مستوى رئيس من الأدنى فالأوسط فالأعلى، والفرق يكمن في النسبة التي يحتويها كل مستوى ثانوي من المستوى الرئيس التالي الأعلى (لا يحتوي الأدنى على أي صفات من المستوى التالي الأعلى، ويحتوي الأوسط على 50 بالمائة تقريباً، ويكون الأعلى معظم الأداء (وليس كله) في المستوى التالي). وقد اختير المحتوى من مناطق مختلفة في الوطن العربي مع التركيز على الجزيرة العربية على اعتبار أن الجامعة الناشئة للاختبارات تقع في الرياض.

ومن أبرز ميزات الاختبار المحوسب أنه يمكن المؤسسة المستخدمة له من أن تطوع الاختبار بحسب حاجتها، سواء من حيث مدة الاختبار، أو مستواه، أو مواضيعه، أو مواطن التركيز (يمكن التركيز على القراءة دون غيرها، مثلاً). وتحدد مستويات المختبرين بناء على المعايير الموضحة في الجداول الآتية:

1.4. احتساب نتيجة اختبار الاستماع (70 سؤالاً) كتقييم وفق معايير المجلس الأمريكي:

جدول 1: معايير احتساب درجة المسموع في اختبار جامعة الملك سعود

التقييم		الرابع 20	الثالث 20	الثاني 15	الأول 15	القسم وعدد أسئلته
أدنى	مبتدئ				10	عدد الإجابات الصحيحة
أوسط					14	
أوسط				4	12	
أعلى				8	14	
أدنى	متوسط			12	14	
أوسط				14	14	
أوسط			4	12	14	
أعلى			16	14	14	
أدنى	متقدم		18	14	14	
أوسط		4	16	14	14	
أوسط		10	18	14	14	
أعلى		16	18	14	14	
---	متميز	18	18	14	14	

2.4. احتساب نتيجة اختبار القراءة (100 سؤال) كتقييم وفق معايير المجلس الأمريكي:

جدول 2: معايير احتساب درجة المقروء في اختبار جامعة الملك سعود

التقييم		الرابع 20	الثالث 20	الثاني 15	الأول 15	القسم وعدد أسئلته
أدنى	مبتدئ				10	عدد الإجابات الصحيحة
أوسط					14	
أوسط				8	12	
أعلى				22	14	
أدنى	متوسط			12	14	
أوسط				22	14	
أوسط			4	18	14	
أعلى			22	22	14	
أدنى	متقدم		18	22	14	
أوسط		4	22	22	14	
أوسط		15	20	20	14	
أعلى		20	22	24	14	
----	متميز	30	24	24	14	

يبين الجدول (1) المتعلق بتقييم مهارة الاستماع، والجدول (2) المتعلق بتقييم مهارة القراءة أن عدد الإجابات في كل مرحلة تحدد المستوى الذي يوضع فيه المختبر، وأن المراحل جميعها تساهم في الحكم النهائي على مستوى المختبر، وليس مرحلة واحدة فقط.

بشكل عام، يظهر العرض السابق أن بعض النقاط المتعلقة ببناء الاختبارات السابقة تحتاج إلى تطوير، فعلى سبيل المثال: نجد أن هذه الاختبارات تركز على المهارات والعناصر والمتطلبات اللغوية ولكنها تهمل عنصراً ضرورياً في الحكم على

الكفاءة اللغوية وهو عنصر الثقافة، وقد وجه النقد نفسه لاختباري التوفل والآيلتس للغة الإنجليزية من قبل عدد من الباحثين (English Help. com, 2016)، كما يظهر أيضا تباينا كبيرا في بناء هذه الاختبارات سيأتي شرحه بالتفصيل فيما يلي.

5. تطوير اختبار قياس الكفاءة اللغوية

يتضح مما سبق أن عملية بناء اختبار قياس الكفاءة اللغوية ليست سهلة، وأنها مازالت في مراحلها الأولى حتى عندما يكون الحديث عن أكثرها شيوعا؛ وهي اختبارات اللغة الإنجليزية التوفل والآيلتس، فعلى الرغم من الشوط الكبير الذي قطعه اللغة الإنجليزية في مجال بناء اختبارات قياس الكفاءة اللغوية، إلا أن الباحثين يشعرون بالقلق حيال هذه الاختبارات، ويرون أن هناك حاجة إلى بناء اختبارات أفضل معتمدة على معايير مدروسة بشكل أكثر علمية، حتى تساهم في تطبيق طرق أفضل في تعليم اللغة الإنجليزية وتعلم للناطقين بغيرها (Free English Help. com, 2016)، والإشكالية الكبرى هنا تكمن في غياب نظرية موحدة ودقيقة لاختبارات اللغات الثانية وتقييمها بشكل عام (Iyldyz, 2007)، إضافة إلى أن اعتماد اختبارات قياس الكفاءة اللغوية على نتائج رقمية مجردة جعل حكمها على المتعلمين غير دقيق، فعلى الرغم من توصل بعض الدراسات إلى نتيجة مفادها أن المتعلمين الذين يحصلون على درجات عالية في اختبار الكفاءة اللغوية، يحصلون أيضا على درجات عالية في مساقات دراساتهم الجامعية فيما بعد، إلا أن معظم الدراسات أثبتت أنه ليس بالضرورة أن يحرز المتعلمون المتفوقون في اختبارات الكفاءة تفوقا في دراساتهم الجامعية (Hyatt & Brooks, 2009). هایت وبروكس)، بل إن كثيرا من تلك الدراسات وجدت أن المتعلمين الذين حصلوا على درجات عالية في اختبار الكفاءة اللغوية، حصلوا على درجات متدنية في دراساتهم الجامعية، وربما يكون أحد التفسيرات لمثل هذه الظاهرة هو الاعتماد الكامل على التوصيف الرقمي لنتائج المختبرين في تحديد كفايتهم اللغوية، وإغفال تحليل

الجوانب المهاريّة والوظيفية للغة، مما يؤدي إلى وضع المختبر في مستوى لا يعكس قدراته اللغوية الحقيقية (Egan, 1999, p. 278).

أما فيما يتعلق باختبارات قياس الكفاءة في اللغة العربية فالأمر أكثر تعقيداً؛ لوجود صعوبات كثيرة تقف في طريق بناء اختبار موثوق وصادق، ومن أهم هذه الصعوبات تشتت الجهود في سبيل بناء الاختبار؛ مما أدى إلى إنتاج نماذج متباينة منه من قبل جهات مختلفة، ويوضح الجدول الآتي أبرز أوجه التباين بين ما تمكن الباحث من حصره من نماذج اختبارات قياس الكفاءة في اللغة العربية حول العالم:

جدول 3: أوجه التباين بين عدد من اختبارات الكفاءة في اللغة العربية

الرقم	الاختبار	مجالات القياس	المرجعية	الفئة المستهدفة	نوع اللغة	النطاق	الوسيلة
1.	العرفان	خمسة مجالات (الكتابة، والقراءة، والاستماع، والتحدث، إضافة إلى مهارة اللغة)	الإطار الأوروبي (CEFR)، برومترك	الناطقون بغيرها (في المستوى المتقدم).	فصيحة	عالمي	الحاسوب، الشبكة
2.	الأكاديمية العربية	خمسة مجالات (فهم المسموع، والقراءة، وبنية الجملة، والكتابة، والمحادثة)	غير معروفة	الناطقون بغيرها	فصيحة	عالمي	الحاسوب (في الاستماع، والقواعد، والكتابة)، شخصي (الكتابة، والمحادثة)

الرقم	الاختبار	مجالات القياس	المرجعية	الفئة المستهدفة	نوع اللغة	النطاق	الوسيلة
3.	الكفاءة في اللغة العربية للناطقين بغيرها، مركز القياس، السعودية	خمسة مجالات (الكتابة، والقراءة، والاستماع، والتحدث، إضافة إلى مهارة اللغة)	غير معروفة	الناطقون بغيرها	فصيحة	عالمي	غير معروف
4.	الكفاءة في اللغة العربية للناطقين بغيرها، الجامعة الإلكترونية، السعودية.	خمسة مجالات (القراءة، والاستماع، والكتابة، والتحدث، واللغة)	غير معروفة	الناطقون بغيرها	فصيحة	غير معروف	غير معروف
5.	جامعة مولانا مالك، إندونيسيا	أربعة مجالات (الاستماع والقراءة والقواعد والكتابة)	غير معروفة	الناطقون بغيرها	فصيحة	محلي	غير معروف
6.	اختبار الكفاءة في اللغة العربية-جامعة ليبينج	أربعة مجالات (الاستماع، والمحادثة، والقراءة، والكتابة)	الإطار الأوروبي (CEFR)	الناطقون بغيرها	العامية	عالمي	غير معروف

الوسيلة	النطاق	نوع اللغة	الفئة المستهدفة	المرجعية	مجالات القياس	الاختبار	الرقم
ورقي	محلي	فصيحة	الناطقون بغيرها	مستويات مسابقات الجامعة نفسها في برنامج اللغة العربية	أربعة مجالات (الاستماع، والتعبير، والترجمة إلى الإنجليزية، والترجمة من الإنجليزية)	اختبار جامعة نيويورك	7.
الحاسوب	محلي	فصيحة	الناطقون بغيرها	مستويات برنامج اللغة العربية الخاص بالجامعة	مجال الفهم والاستيعاب	اختبار بيرلنك سوفتوير	8.
الحاسوب	عالمي	فصيحة	الناطقون بغيرها	غير معروفة	ثلاثة مجالات (القراءة، والكتابة، والقواعد)	اختبار سبرانش كافي لانجوجز بلس	9.
ورقي	عالمي	فصيحة	الناطقون بغيرها	الإطار الأوروبي (CEFR)	أربعة مجالات (الاستماع، والقراءة، والكتابة، والمحادثة)	امتحان الكفاءة العالمي للغة العربية للناطقين بغيرها - جامعة النجاح الوطنية - فلسطين	10

الرقم	الاختبار	مجالات القياس	المرجعية	الفئة المستهدفة	نوع اللغة	النطاق	الوسيلة
11.	اختبار الجامعة الأمريكية في القاهرة	غير معروف	غير معروفة	الناطقون بغيرها، والناطقون بها	فصيحة وعامية	غير معروف	غير معروف
12.	اختبار الكفاءة في الجامعة الأردنية	مجالان (القراءة والكتابة)	غير معروفة	الناطقون بغيرها والناطقون بها	فصيحة	محلي ورقني	
13.	اختبار التنازل	ثلاثة مجالات (الوظيفي، المنطوق: استماع وخطاب، المكتوب: استماع، وكتابة، وقراءة، ومحادثة)	غير معروفة	الناطقون بغيرها، والناطقون بها	فصيحة	عالمي	الحاسوب
14.	اختبار العين لقياس الكفاءة اللغوية	مجالان (الشفهي: استماعاً وتحديثاً، و التحريري: قراءة وكتابة)	وثيقة كفايات الاختبار بعد الاستئناس بأشهر المعايير العالمية المعتمدة في تعليم اللغات.	الناطقون بها	فصيحة	محلي	الحاسوب أو ورقني

الوسيلة	النطاق	نوع اللغة	الفئة المستهدفة	المرجعية	مجالات القياس	الاختبار	الرقم
الحاسوب أو ورقي	عالمي	فصيحة	الناطقون بغيرها	معياري المجلس الأمريكي لتعليم اللغات الأجنبية (ACTFL)	أربعة مجالات (الاستماع، والقراءة، والكتابة، والمحادثة)	اختبار جامعة الملك سعود	15.

ويمكن حصر أوجه التباين في عدد من المحاور التي يبينها الجدول أعلاه، وأبرزها:

1.5. مجالات القياس: ويقصد بها ما يقيسه الاختبار من مهارات اللغة أو عناصرها أو متطلباتها، ويلاحظ أن عدد المجالات التي قاستها هذه الاختبارات تراوح بين مجالين وخمسة مجالات؛ كما يلاحظ أيضاً أن هذه المجالات اختلفت من اختبار إلى آخر فبعضها قسمها بحسب المهارات والعناصر (العرفان، الأكاديمية العربية، على سبيل المثال)، وبعضها قسمها بحسب المهارات والوظائف (جامعة نيويورك)، وبعضها قسمها بحسب المهارات فقط (جامعة الملك سعود)، وبعضها جمع مهارتين في مجال واحد، مثل: الاستماع والتحدث في المجال الشفهي (العين).

2.5. المرجعية: ويتضح هنا أن بعض هذه الاختبارات استند إلى مرجعية معينة، مثل: معايير المجلس الأمريكي للغات (ACTFL)، أو الإطار الأوروبي (CEFR)، مثل: اختبار العرفان، واختبار جامعة الملك سعود، واختبار جامعة النجاح، وبعضها طور مرجعية خاصة به معتمداً على المرجعيات العالمية، مثل: اختبار العين، وبعضها اعتمد على تصنيفه الداخلي لمستويات البرنامج الخاص به، مثل: جامعة نيويورك، وبعضها لم يفصح عن المرجعية التي استند إليها في بناء الاختبار.

وتجدر الإشارة هنا إلى أن تلك الجهات التي اعتمدت على توصيف مستوياتها أو مساقاتها التي تدرس ضمن برنامجها للغة العربية في بناء الاختبار، خالفت واحداً

من خواص اختبار قياس الكفاءة اللغوية، ألا وهو عدم ارتباط هذا النوع من الاختبارات بالمقررات الدراسية لمؤسسة ما، وربما يكون هذا هو السبب الذي قاد مؤسسات أخرى إلى عدم الاعتراف بهذا الاختبار لقياس مستوى كفاءة طلبتها، كما حدث مع اختبار جامعة نيويورك الذي لم تعترف به بعض الجامعات الأمريكية.

3.5. الفئة المستهدفة: صممت بعض هذه الاختبارات لقياس كفاءة غير الناطقين باللغة العربية (العرفان، وجامعة الملك سعود على سبيل المثال)، وبعضها لقياس كفاءة الناطقين بها (العين على سبيل المثال)، وبعضها لقياس كفاءة الناطقين وغير الناطقين (الجامعة الأمريكية في القاهرة)، والنوع الثالث الذي يقيس كفاءة الناطقين وغير الناطقين بالعربية يبقى موضع تساؤلات كثيرة، تدور حول جواز قياس فئتين مختلفتين باستخدام الاختبار نفسه، وهل هناك جوانب تخص فئة أخرى دون غيرها قد يغفلها هذا الاختبار، مثل: النطق، أو تركيب الجملة على سبيل المثال لا الحصر؟

4.5. نوع اللغة: قاست معظم الاختبارات التي نوقشت في هذا البحث كفاءة المختبرين في مهارات اللغة العربية الفصيحة وعناصرها ومتطلباتها، إلا أن واحدا منها صمم لقياس الكفاءة في العامية (اختبار جامعة ليبزج)، مقسما الاختبار بحسب المناطق في العالم العربي، وقاس اختبار آخر الفصيحة والعامية (اختبار الجامعة الأمريكية في القاهرة).

5.5. النطاق: ادعى القائمون على عدد كبير من هذه الاختبارات أنها اختبارات علمية ويمكن اعتمادها من قبل الجامعات والمؤسسات والهيئات المهتمة على اختلافها، ووضح عدد قليل من القائمين على هذه الاختبارات بأنها اختبارات محلية؛ أي أنها تستخدم في المؤسسة التي صممتها فقط، وهو أمر ربما يتنافى وطبيعة استخدامات اختبارات قياس الكفاءة اللغوية المتعارف عليها، وهي: العالمية. على الرغم من ذلك فرمما يكون الاختبار الوحيد من بين الاختبارات الواردة في الجدول (3) الذي اعتمد من قبل مؤسسة غير المؤسسة التي صممتها، هو: اختبار جامعة النجاح.

6.5. الوسيلة: ويقصد بها استخدام الحاسوب أو النموذج الورقي في تنفيذ الاختبار، ويظهر من الجدول (3) أن بعض الاختبارات اعتمدت كلياً على الحاسوب سواء عبر الشبكة أو في المختبرات (العرفان على سبيل المثال)، وبعضها جمع بين الحاسوب والنموذج الورقي (الأكاديمية العربية على سبيل المثال)، وبعضها اكتفى بالنموذج الورقي التقليدي (اختبار جامعة نيويورك على سبيل المثال).

إضافة إلى ما سبق، فإنه يمكن تحديد وجه آخر من أوجه التباين بين هذه الاختبارات، وهو النسب المعتمدة لتقييم المهارات، ويبين الجدول الآتي هذه النسب لبعض هذه الاختبارات، ويبقى عدد منها لم يوضح هذه النسب في أدياته:

جدول 4: نسب تقييم المهارات في اختبارات الكفاءة في اللغة العربية

الرقم	الاختبار	الاستماع	القراءة	الكتابة	المحادثة
1.	العرفان	غير معروف	غير معروف	غير معروف	غير معروف
2.	الأكاديمية العربية	غير معروف	غير معروف	غير معروف	غير معروف
3.	الكفاءة في اللغة العربية للناطقين بغيرها، مركز القياس، السعودية	25 %	40 %	35 %	لا يوجد
4.	الكفاءة في اللغة العربية للناطقين بغيرها، الجامعة الإلكترونية، السعودية.	غير معروف	غير معروف	غير معروف	غير معروف
5.	جامعة مولانا مالك، إندونيسيا	غير معروف	غير معروف	غير معروف	غير معروف
6.	اختبار الكفاءة في اللغة العربية - جامعة ليبينج -	25 %	25 %	25 %	25 %
7.	اختبار جامعة نيويورك	لا يوجد	غير معروف	غير معروف	لا يوجد
8.	اختبار بيرلنك سوفتوير	لا يوجد	غير معروف	غير معروف	لا يوجد
9.	اختبار سبراتش كافي لانجوجز بلس	لا يوجد	غير معروف	غير معروف	لا يوجد

الرقم	الاختبار	الاستماع	القراءة	الكتابة	المحادثة
10.	امتحان الكفاءة العالمي للغة العربية للناطقين بغيرها - جامعة النجاح الوطنية - فلسطين	25 %	25 %	25 %	25 %
11.	اختبار الجامعة الامريكية في القاهرة	غير معروف	غير معروف	غير معروف	غير معروف
12.	اختبار الكفاءة في الجامعة الأردنية	لا يوجد	غير معروف	غير معروف	لا يوجد
13.	اختبار التتال	غير معروف	غير معروف	غير معروف	غير معروف
14.	اختبار العين لقياس الكفاءة اللغوية	غير معروف	غير معروف	غير معروف	غير معروف
15.	اختبار جامعة الملك سعود	45 % من عدد الأسئلة	45 % من عدد الأسئلة	7 % من عدد الأسئلة	منفصل عن بقية الاختبار، ويحدد المقيم مستوى المختبر فوراً

يوضح الجدول (4) تباينا في نسب تقييم المهارات اللغوية بين الاختبارات، ففي حين أغفلت بعض الاختبارات مهارة أو أكثر، ولم تقسها أساسا (على الأغلب الاستماع والمحادثة، اختبار بيرلنك على سبيل المثال)، نجد اختبارات أخرى قاست المهارات جميعها، إلا أنها اختلفت في توزيع نسب التقييم؛ ففي حين ساوى اختبار جامعة ليبينج بين نسب تقييم المهارات وأعطى نسبة 25 % لها جميعا، أعطى اختبار مركز القياس في السعودية نسبا مختلفة لثلاث مهارات؛ بحيث أفرد نسبة 25 % للاستماع، ونسبة 40 % للقراءة، و35 % للكتابة، ولم يقس مهارة المحادثة.

وعلى الرغم من اعتماد الاختبارات السابقة على الدرجات الممنوحة لكل مهارة في نسب التقييم، اعتمدت اختبارات أخرى على نسبة عدد الأسئلة لكل مهارة وفي كل مستوى، ويتضح ذلك في اختبار جامعة الملك سعود الذي أفرد 45% من الأسئلة

لمهارة الاستماع مقسمة على تسعة مستويات، و45% مهارة القراءة مقسمة أيضا على تسعة مستويات، و10% مهارة الكتابة مقسمة كذلك على تسعة مستويات، فيما فصل مهارة المحادثة عنها جميعا وجعل تقييمها خاصا يحدد فيه المقيم مستوى المختبر بناء على جدول معايير (Rubric) خاص بهذه المهارة يوضح أداء المختبر المتوقع في كل مستوى، بحسب معيار المجلس الأمريكي للغات.

يظهر بوضوح من النقاش السابق أننا لا نفتقر إلى اختبارات قياس الكفاءة في اللغة العربية سواء للناطقين بها أو لغير الناطقين بها من حيث العدد، إلا أن المؤسسات والهيئات المهتمة بمجال تعليم العربية لم تعتمد واحدا منها اختبارا عالميا، فكل مؤسسة تستخدم اختبارها فقط ولم تتمكن من تحويله إلى اختبار عالمي، ومؤسسات أخرى لم تكثرث إلى تطوير اختبار خاص بها أساسا؛ لأنها ربما لا تدرك وجود هذه الاختبارات أصلا، وتتذمر من غيابها؛ لذلك فالمشكلة ليست في تطوير واستخدام هذا النوع من الاختبارات، وإنما في التباين الواضح في بنائها، وتطبيقاتها، وإجراءاتها، وشيوعها، كما ظهر سابقا، وهو دليل على تخط، ربما يضعف الثقة في هذه الاختبارات ويكون سببا في قلة استخدامها، وضيق انتشارها. وربما يكون وراء ذلك غياب الجهة الراعية التي تكفل تنسيقا فاعلا بين الجهات المهتمة والمستخدمة لهذه الاختبارات، وغياب الأسس النظرية والعملية لبناء هذه الاختبارات، والتي من أبرزها الإطار المرجعي الموحد المبني على أسس علمية مستندة إلى حاجة الميدان، وما توصلت إليه نظريات اكتساب وتقييم اللغة الثانية.

6. مصداقية أدوات قياس الكفاءة اللغوية في مهارة المحادثة

في دراسة أجراها الشاطر (Alshatter, 2017) حول مصداقية أدوات قياس الكفاءة اللغوية في مهارة المحادثة، اختبر مصداقية تقييم عشرة مقيمين لمحدثات مسجلة لخمسة متعلمين، وذلك بإجراء التقييم ثلاث مرات للمحدثات المسجلة نفسها، للمتعلمين أنفسهم. ترك الباحث مدة ستة أسابيع بين كل تقييم وآخر، وفي كل مرة كان يطلب من المقيم أن يستخدم معيارا مختلفا لتقييم المتعلمين؛ ففي المرة الأولى

طُلب من المقيمين تطبيق المعيار نفسه الذي يستخدمونه في اختباراتهم العادية، وفي المرة الثانية زود الباحث المقيمين بمعيار أعد بالاستناد إلى أربعة معايير أمريكية لاختبار المهارات الشفهية، هي: اختبار مديرية تعليم ماساتشوست لتقييم المهارات الأساسية (Massachusetts) Department of Education Assessment of Basic Skills)، واختبار فيرفكس كاونتي لمدارس ولاية فرجينيا الحكومية (Fairfax County)، واختبار مدارس أوغدن الإعدادية (Virginia Public schools)، واختبار مدارس أوغدن الإعدادية (Ogden Elementary School)، واختبار مقاطعة بلدة إلينوي 214 (Illinois Township District 214)، وهي جميعها اختبارات تقيس مهارة المحادثة باعتماد معايير مختلفة قام الباحث بتوحيدها في معيار واحد، إلا أن العامل المشترك بينها هو أنها جميعها تعتمد على ما يسمى بجداول الأداء (Rubrics) في تقييم المختبرين وتحديد مستوياتهم، وتتسم بنود هذه الجداول بأنها انطباعية وتعتمد على تقديرات المقيمين، أما المرة الثالثة فقد زود الباحث المقيمين بمعيار قياس طوره بنفسه معتمدا على معايير القياس سابقة الذكر إضافة إلى عدد من المعايير الأخرى جاءت في دراسات مختلفة تتحدث عن اختبارات قياس المهارات الشفهية، وقد تميز المعيار الذي استخدم في المرة الثالثة بأنه ينقسم إلى قسمين: قسم يعتمد على الإحصاء في التقييم، وآخر يعتمد على انطباع المقيم المقنن.

أكدت نتائج الدراسة أن تقييم المقيمين المختلفين للمختبر نفسه يختلف باختلاف المعيار المستخدم في التقييم، كما أظهرت أن تقييم المقيم الواحد للمختبر نفسه في عدة مرات يختلف باختلاف المعيار المستخدم، وأظهرت أيضا أن استخدام معيار موحد يزيد من مصداقية الاختبار، كما ظهر من نتائج الدراسة أن المعايير التي تعتمد على بنود تقاس إحصائيا أكثر دقة ومصداقية من تلك التي تعتمد على بنود تقاس بالانطباع المقنن. وقد بينت الدراسة أن تطوير أداة لقياس المهارات اللغوية الشفهية ليست بالمهمة السهلة لعدة أسباب، من أبرزها:

(1) لوازم إدارة هذا النوع من الاختبارات،

(2) طبيعة مهارات المحادثة،

(3) حاجات المتعلمين،

(4) أهداف المنهج،

(5) قدرة المقيمين وتدريبهم، والأهم من ذلك كله،

(6) بناء أداة القياس نفسها.

7. دور التقنية (التكنولوجيا) في تطوير اختبارات قياس الكفاءة اللغوية

هناك شبه إجماع على أن استخدام التقنيات الحديثة بشكل عام، واستخدام الحاسوب على وجه الخصوص يسهل عملية تقويم مهارات اللغة وما يرتبط بها من عناصر ومتطلبات، ويشمل ذلك جوانب اللغة كلها، من استماع وقراءة ومحادثة وكتابة وقواعد وثقافة وأدب وبلاغة، وغيرها. وقد نجح عدد كبير من الاختبارات التي تناولتها هذه الورقة بالبحث في توظيف تقنية الحاسب في تقييم مهارات اللغة العربية وبالذات مهارتي الاستماع والقراءة، كما نجحت أيضا في تقييم عناصرها وبالأخص عنصرى المفردات والقواعد، وربما من السهل أيضا تطويع هذه التقنية لتقييم مهارة الكتابة، أما فيما يتعلق بقياس مهارة المحادثة وما يمكن أن تقدمه التقنيات الحديثة للمساهمة في الوصول إلى تقييم موضوعي صادق ودقيق، فإن الوضع يبدو أكثر صعوبة مقارنة بقياس المهارات الأخرى سابقة الذكر، خاصة عندما يكون الحديث عن اختبار عدد كبير من المختبرين (Kitao & Kitao, 1996)، كيتاو و كيتاو، ص 2)، حتى لو تمكنا من تطوير هذا النوع من الأدوات، فإن إدارة الاختبار نفسه تبقى مهمة صعبة (Rubin, Daly, Mc Croskey, & Mead, 1982)، روبن وآخرون، ص 117). وهنا يأتي دور التقنية التي يمكن أن تشكل جزءا رئيسا من الحل؛ حيث أظهر استخدام تقنية الحاسب في اللغات تسارعا في الانتشار في الآونة الأخيرة بما يسمى بعون الحاسب في تعليم اللغات (Computer Assisted Language Learning CALL)، والتي - التقنية - مكنت المتعلمين من الانخراط في عدد من المهارات اللغوية، مثل: الاستماع، والقراءة، وأنشطتها، مثل: ملء الفراغات

(Egan, إيجان 1999، ص 277).

يمكن النظر إلى برامج عون الحاسب هذه على أنها عنصر واعد وضروري في زيادة دقة ومصداقية وسهولة اختبارات تقييم المهارات الشفهية، خاصة أنه من المتوقع أنها أسرع وأدق في قياس المهارات اللغوية وبالذات مهارة المحادثة من الطرق التقليدية في القياس، وأكثرها شهرة في الوقت الحاضر فيما يتعلق بالمحادثة ما يدعى بتقنية التعرف الآلي على الأصوات (Automatic Speech Recognition-ASR)، والتي تشكل خطوة إلى الأمام في طريق تطوير اختبارات آلية لمهارة المحادثة؛ حيث أظهرت بعض النماذج نسبة أقل من 4% من الاختلاف بين نتائج التقييم لهذه النتائج، وتقييم المقيمين من البشر (Johnson, جونسون 2012، ص 39)؛ لذلك فإنه من المتوقع أن تحل هذه التقنية عددا كبيرا من مشاكل اختبار المهارات الشفهية، مثل: التكلفة العالية، والموارد البشرية اللازمة، والجهود الإدارية المطلوبة. وتجدر الإشارة إلى أن تقنية العون الحاسوبي هذه قدمت حتى الآن خدمات كثيرة لتعليم اللغات باستخدام برامج التعرف الآلي على الحديث، منها:

(1) التعرف على جنسيات المتحدثين بناء على أدائهم اللغوي الشفهي (Isulaiman et al., 2012، السليمان وآخرون، ص 107)،

(2) تقييم جودة النطق لدى متعلمي اللغة الثانية وبدقة عالية قريبة جدا من تقييم المقيمين من البشر،

(3) التعرف على نوع الأخطاء في النطق وتقييمها،

(4) التقاط أخطاء النطق المتعلقة بالطلاقة،

(5) تزويد المتعلمين بالتغذية الراجعة لتصويب أخطائهم (وت، Witt، 2012، ص 2-3).

وعليه فإنه إن كان بالإمكان استخدام برامج العون الحاسوبي لتقديم كل هذه الخدمات في تقييم النطق، فإنه من المتوقع أيضا أن نتمكن من استخدامها في قياس أداء المتعلمين اللغوي الشفهي، وتزويدهم بتغذية راجعة مفيدة (Batliner, Honig،

(Noth, 2012 & هونج، باتلنز، نوث، ص 29)، إضافة إلى ذلك ستمكن هذه البرامج المتعلمين من لعب دور بارز في تعلمهم للغة وتعلم الآخرين لها، وذلك عبر تقييم مستمر واضح وثابت لمهاراتهم الشفهية، والتي يمكن أن تكون جزءاً من أنشطة التعلم المستمر لدى هؤلاء المتعلمين (Boud، بود، 2000، ص 152)، ولكن هذا الهدف المهم لاختبار المهارات الشفهية أو الأداء التواصلي آلياً لا يمكن أن يتحقق إلا بتحقيق عدد من العناصر في تعليم اللغات، والتي من أبرزها وضع معايير لبناء المحتويات ومعايير للتقويم، وهنا تبرز أهمية التركيز على تطوير معايير إحصائية لتقييم المهارات الشفهية؛ وذلك لأن تلك المعايير أكثر قابلية للحوسبة من المعايير الانطباعية.

خاتمة

تخلص هذه الورقة إلى أن بناء اختبار يقيس الكفاءة اللغوية لدى متعلمي العربية من غير الناطقين بها ليس أمراً سهلاً، وأنه وعلى الرغم من توافر عدد لا بأس به من هذه الاختبارات سواء في العالم العربي أو خارجه، إلا أن تشتت الجهود، وغياب إطار مرجعي يعتد به في بناء هذا الاختبار، وضعف التنسيق بين الجهات المهتمة، جعل من الصعب بناء اختبار عالمي موضوعي ودقيق وصادق؛ لذلك فإن عدداً من الخطوات يجب اتخاذها من أجل الوصول إلى بناء هكذا اختبار. وتجدر الإشارة إلى أن استخدام التقنية الحديثة وعلى الأخص الحاسوب سيجعل من قياس الكفاءة اللغوية مهمة أسهل، وأدق، وأقل كلفة، خاصة مع الانتشار السريع لهذه البرامج، وكثرة وسائل تطوير هذه التقنية، وزيادة اهتمام المختصين بها، وظهور عدد من الدراسات التجريبية التي أثبتت نجاحها ومصداقيتها. وعليه، فإنه ومن أجل التوصل إلى بناء اختبار عالمي لقياس الكفاءة اللغوية في اللغة العربية، فإنه يلزمنا أن نضع في عين الاعتبار عدداً من المتطلبات التي ستساهم في بناء اختبار دقيق، وموضوعي، وذو مصداقية وثبات، ومن أبرزها:

1- إنشاء مؤسسة خاصة ببناء اختبار قياس الكفاءة اللغوية في اللغة العربية،

تتبنها واحدة من المؤسسات المحورية في العالمين العربي والإسلامي، مثل: منظمة العالم الإسلامي، أو جامعة الدول العربية، أو ما شابههما تكون مسؤولة عن بناء الاختبار وإدارته،

2- التمييز بين أنواع الاختبارات المختلفة (كفاءة، وتحديد مستوى، وتشخيصي، الخ)، وبناء اختبار الكفاءة بعد اعتماد التعريف الدقيق لهذا النوع من الاختبارات،

3- تطوير إطار مرجعي عام قبل الشروع في بناء أي اختبار يحدد مواصفات الاختبار، ويوضح المخرجات المتوقعة من المختبرين في كل مستوى للمهارات والعناصر اللغوية، والمتطلبات اللغوية جميعها،

4- وضع أسس واضحة ودقيقة للتقييم، تبين الإجراءات التي يجب اتباعها في كل مرحلة من مراحلها، تضمن مصداقيته وثباته،

5- العمل الجاد من أجل تطوير نموذج محوسب من الاختبار يكون التقويم فيه آلياً.

6- ضرورة تكوين لجان بناء الاختبار من مختصين في اللغة العربية، ومختصين في اللغويات التطبيقية، ومختصين في القياس والتقويم، ومختصين في تقنيات تعليم اللغات وتقييمها،

7- تضافر الجهود بين مختصين في تطوير اختبارات اللغة العربية، وآخرين مختصين بتطوير اختبارات لغات أخرى.

قائمة المراجع

باللغة العربية:

- التنال، (2015). اختبار شهادة الكفاءة الدولية في اللغة العربية.
Retrieved from <http://www.attanalalarabi.com/>
- الجامعة الأردنية. (2016). امتحان الكفاية باللغة العربية لطلبة الدراسات العليا.
Retrieved from <http://graduatedstudies.ju.edu.jo/ar/arabic/ArabicTest.aspx>
- الجامعة الأمريكية بالقاهرة. (2016). الاختبارات التقويمية والتقييم (للأفراد/المؤسسات).
Retrieved from <http://www.aucegypt.edu/ar/academics/continuing-education>.
- الجامعة السعودية الإلكترونية. (2017). اختبار اللغة العربية المعياري لغير الناطقين بها.
Retrieved from https://www.seu.edu.sa/sites/ar/colleges/CSTS/Arabic_nn%20Department/Pages/About.aspx.
- العرفان. (2014). اختبار اللغة العربية للناطقين بغيرها.
Retrieved from <http://e-lang.org/ar/elang-test.php>.
- المركز الوطني للقياس. (2014). الاختبارات اللغوية: اختبار اللغة العربية المقنن لغير الناطقين بها.
Retrieved from <http://www.qiyas.sa/Tests/LanguageTests/Pages/Arabic.aspx>.
- جامعة الإمارات العربية المتحدة. (2012). اختبار العين لقياس الكفاءة اللغوية.
Retrieved from <http://alain-proficiency-center.blogspot.com/p/blog-page.html>
- جامعة النجاح. (2014). امتحان الكفاءة العالمي للغة العربية للناطقين بغيرها.
- جامعة مولانا مالك. (2015). اختبار اللغة العربية الموحد.

Retrieved from <http://ppba.uin-malang.ac.id/ar/>.

- عبد الخالق، م. (1996). اختبارات اللغة ط 2. جامعة الملك سعود الرياض،
السعودية.

باللغة الأجنبية:

- Al-Arabiyya. (2016). Arabic Language Proficiency Test. Retrieved from <http://www.test-arabic.com/en/home>
- Alshatter, G. H. (2017). Validity of Oral Tasks Testing in Arabic L2 Teaching. *LINGUISTICA COMMUNICATIO*, 18(1-2), 16.
- Alsulaiman, M., Mohamed, B., Al Shatter, G., Al-Kahtani, S., Muhammad, G., Ali, Z., & Al-Gabri, M. (2012). "Automatic identification of Arabic L2 learners origin". *International Symposium on Automatic Detection of Errors in Pronunciation Training*, 107-112.
- Arab-Academy. (2015). Arabic Language Proficiency Test. Retrieved from <http://www.arabacademy.com/features/arabic-language-proficiency-test/>.
- AZU. (2007). Types of Language Tests. *Lesson Notes, Teaching Resources and Research in Applied Linguistics & ESL*. Retrieved from <http://www.azlifa.com/types-of-language-tests/>
- Bellevue. (2016). New York University (NYU) Foreign Language Proficiency Exams. Retrieved from <http://www.bellevue.edu/student-support/testing-center/nyu-foreign-language-proficiency-exams>.
- Boud, D. (2000). "Sustainable Assessment: Rethinking Assessment for the Learning Society". *Studies in continuing education*, 22, 16.
- Egan, K. B. (1999). "Speaking: A critical skill and challenge". *CALICO*

- Journal*, 16, 277-293.
- Ezekiel, L. L. T. R. (2016). Types of Tests: Learn How to Succeed on Different Types of Exams. Retrieved from <http://www.goodluckexams.com/types-of-tests/>
 - Free English Help.com. (2016). Pros and Cons of International Proficiency Tests Such as IELTS and TOEFL Retrieved from <https://freeenglish-help.com/international-proficiency-english-language-tests>
 - Honig, F., Batliner, A., & Noth, E. (2012). "Automatic Assessment of Non-Native Prosody - Annotation", Modelling and Evaluation. *International Symposium on Automatic Detection of Errors in Pronunciation Training*, 21-30.
 - Hyatt, D., & Brooks, G. (2009). Investigating Stakeholders' Perceptions of IELTS as an entry requirement for higher education in the UK. *Project report for British council/Cambridge ESOL/IDP: IELTS Australia - UK1120 (round 12, 2007)*.
 - Iyldyz, L. B. (2007). "Rethinking Validity of the L2 Proficiency Concept: Lessons for EIL". *The Asian EFL Journal Quarterly*, 9(1), 65-96.
 - Johnson, W. L. (2012). "Error detection for teaching communicative competence". *International Symposium on Automatic Detection of Errors in Pronunciation Training*, 3742.
 - Kitao, S. K., & Kitao, K. (1996). Testing speaking, 1-7.
 - PearlinkSoftware. (2016). Arabic Language Proficiency Test. Retrieved from <http://www.earabicpro.com/products/proficiencyTest.htm>
 - Rubin, D. L., Daly, J., McCroskey, J. C., & Mead, N. A. (1982, October,

1982). "A review and critique of procedures for assessing speaking and listening skills among preschool through grade twelve students". *Communication education*, 31, 18.

- Sprachcaffe. (2016). Online Arabic Proficiency Test. Retrieved from <http://www.sprachcaffe.com/english/language-test/arabic-test-welcome.htm>
- Witt, S. M. (2012). Automatic error detection in pronunciation training: where we are and where we need to go. International Symposium on Automatic Detection of Errors in Pronunciation Training, 1-8.