ATATÜRK ÜNİVERSİTESİ YAYINLARI
ATATURK UNIVERSITY PUBLICATIONS

Research Article *Araştırma Makalesi* DOI: 10.5152/ilted.2022.221540

بناء وتطوير مُدوَّنة لُغويَّة مُتوازية بينَ التُّركيَّة والعربيَّة لأغراض تعليم اللُّغة العربيَّة للنَّاطقينَ بالتُّركيَّة

Ana Dili Türkçe Olanlara Arapça Öğretmek Amacıyla Türkçe ile Arapça Arasında Paralel Bir Dil Derlemi Oluşturma ve Geliştirmeye Yönelik Bir İnceleme

Building and Developing a Parallel Corpus between Turkish and Arabic for The Purposes of Teaching Arabic to Turkish Native Speakers

Almoataz B. AL-SAID

Cairo University, Faculty of Dar Al Uloom, Department of Linguistics, Egypt

الملخص

تقومُ فكرةُ هذه الدِّراسة على الرَّبط بينَ القُدرة على اكتساب اللُّغة الثَّانية وطبيعة البيئة الَّتي تُستَخدَمُ فيها هذه اللُّغة. وينشُدُ الباحثُ تطبيقَ فكرتِه على نموذج تعليم العربيَّة للنَّاطقينَ بالتُّركيَّة، نظرًا لحاجة هؤلاء إلى استخدام العربيَّة في مُجتمعهم التُّركيّ. ولأنَّ تعليمَ اللُّغة العربيَّة يستدعي الوُقُوفَ على واقع استعمالِها، سواءٌ أكانَ هذا الاستعمالُ في بيئتها العربيَّة أم في بيئةٍ أخرى، فإنَّ هذه الدِّراسة تسعى إلى تقديم رؤيةٍ منهجيَّة لكيفيَّة بناء مُدوَّنةٍ لُغويَّةٍ مُتوازية [ثُنائيَّة اللُّغة] بينَ: التُّركيَّة والعربيَّة، ثُمَّ مُعالجة هذه المُدوَّنة لاستخلاص المعارف الرَّئيسة من المُفرَدات والتَّراكيب [العربيَّة] الأكثر مُناسبَةً لبيئة المُتعلِّم [التُّركيَّة]، استرشادًا بمعايير الإطار الأوروبيّ المرجعيّ العامّ للُّغات (CEFR). وقد التزَمَت الدِّراسةُ المنهجَ الوصفيَّ الَّذي يُعنى بدراسة الظَّواهر اللُّغويَّة والتَّعبير عنها كمًّا وكيفًا، على الصُّورة الَّتي تبدو عليها اللُّغةُ في الواقع، انطلاقًا من أدبيَّات لسانيَّات المُدوَّنة من ناحيةٍ واللِّسانيَّات التَّربويَّة من ناحيةٍ أخرى. وتَهدفُ الدِّراسةُ إلى مُعالجة الفجوة بينَ المُقرَّرات التَّعليميَّة للعربيَّة – بمُستوياتِها المُختلفة – والبيئة [غير العربيَّة] الَّتي يعيشُ فيها المُتعلِّم التُّركيّ؛ حيثُ يَجِدُ هذا المُتعلمُ نفسَه مُضطرًّا إلى تعلُّم كثيرٍ من المُفرَدات والأساليب والتَّراكيب الَّتي قد تكونُ غريبةً أو نادرة الاستخدام في مُجتمَعه، الأمر الَّذي ينعكسُ على قُدرته على مُمارسة اللُّغة والتَّمكُّن منها. وقد توصَّلَت الدِّراسةُ إلى جُملةٍ من النَّتائج، منها: تطوير مُدوَّنة لُغويَّة مُتوازية بينَ التُّركيَّة والعربيَّة ومُعالجة مادَّتها لغاية تعليم العربيَّة للنَّاطقينَ بالتُّركيَّة، وتقديم تصوُّر منهجيّ لكيفيَّة توظيف هذه المُدوَّنة في تعليم العربيَّة كلُغة ثانية، وطرح مُقاربة اقتصاديَّة تُساعدُ على الإفادة من الموارد التَّعليميَّة في توفير الوقت والجُهد المَبذُولَين في عمليَّتَي التَّعليم والتَّعلُّم.

الكلمات المفتاحيَّة: مُدوَّنة لُغويَّة مُتوازية، علم اللُّغة التَّربويّ، تعليم العربيَّة للنَّاطقينَ بغيرِها، اللُّغة التُّركيَّة.

ÖZ

Bu çalışma, ikinci dili edinme yeteneği ile dilin kullanıldığı toplumun doğası arasında bağlantı kurmayı hedeflemektedir. Araştırmacı, fikrini, Türk toplumunda Arapça kullanım ihtiyaçları göz önünde bulundurularak, ana dili Türkçe olanlara Arapça öğretimi örnekliği üzerinden uygulamaya çalışmaktadır. Arapçayı öğretmek, ister Arap toplumunda, isterse başka bir toplumda olsun, kullanımının gerçekliği üzerinde durmayı gerektirdiğinden, bu çalışma Türkçe ve Arapça arasında iki dilli paralel bir derlemin nasıl kurulacağına ve nasıl işleneceğine dair sistematik bir vizyon sağlamayı, daha sonra Avrupa Ortak Dil Referans Çerçevesi kriterleri (CEFR) rehberliğinde, Arapça öğrenen Türk toplumuna en uygun temel Arapça kelime ve dil yapılarını ortaya koyup, bu derleme işlemeyi amaçlamaktadır. Çalışma, bir yandan dil bilimin yazılı literatürüne, diğer yandan eğitim dil bilimine dayalı olarak, dilin gerçekte göründüğü forma göre, dilsel olguları incelemek ve bunları nicel ve nitel olarak ifade etmekle ilgilenen betimleyici yaklaşıma bağlı kalmıştır. Çalışma, çeşitli seviyelerdeki Arapça eğitim müfredatı ile Türk öğrencinin yaşadığı [Arap olmayan] ortam arasındaki mesafeyi ele almayı hedeflemektedir. Şöyle ki; öğrenci kendini, bulunduğu toplumda garip görülen veya nadiren kullanılan birçok kelime, üslup ve yapıyı öğrenmek zorunda hissetmekte, bu ise, öğrendiği dilde pratik yapma ve ustalaşma yeteneğini olumsuz bir şekilde etkilemektedir. Araştırma, Türkçe ve Arapça arasında paralel dilsel bir derlemin geliştirilmesi ve materyallerin ana dili Türkçe olanlara Arapça öğretmek amacıyla işlenmesinin gerekliliği, ikinci dil olarak Arapça öğretiminde nasıl kullanılacağına dair sistematik bir tasavvur sunulması, bu derlemin eğitim-öğretim sürecinde harcanan zaman ve emekten tasarruf etmede, eğitim kaynaklarından yararlanmaya yardımcı olan ekonomik bir yaklaşım ortaya koyması gibi sonuçlara ulaşmıştır.

Anahtar Kelimeler: Paralel Dil Derlemi, Eğitim Dil Bilimi, Ana Dili Arapça Olmayanlara Arapça Öğretimi, Türkçe.

ABSTRACT

The idea of this study is to try to connect between the ability to acquire the second language and the society's nature in which the language is used. The researcher seeks to apply his idea on the model of teaching Arabic to Turkish native speakers, in view of their need to use Arabic in their Turkish society. Because teaching Arabic language requires understanding the reality of its usage, whether that usage in its Arab society or in another society, this study seeks to provide a methodological vision of how to build a bilingual parallel corpus between Turkish and Arabic, then process this corpus to extract the the most suitable knowledge of Arabic vocabulary and structures adapted to Turkish learner society. The study will be guided by the Common European Framework of Reference for Languages criteria. This research is based on a descriptive approach studying language

Received/Geliş Tarihi: 07.02.2022
Accepted/Kabul Tarihi: 10.03.2022

Sorumlu Yazar/Corresponding Author: Almoataz B. AL-SAID
E-posta: moataz@cu.edu.eg

Cite this article: Al-Said, Almoataz. "Building and Developing a Parallel Corpus Between Turkish and Arabic for the Purposes of Teaching Arabic to Turkish Native Speakers". *Journal of Ilahiyat Researches* 57/1 (June 2022), 18-28.

Atıf: Al-Said, Almoataz. "Binā wa Tatvīr Mudavvana Lugaviyye Mutavāziye Bayna al-Turkiyya wa al-Arabiyya li Agrādi Talīm al-Luga al-Arabiyya li al-Nātiqīna bi al-Turkiyya". *İlahiyat Tetkikleri Dergisi* 57/1 (Haziran 2022), 18-28.

Content of this journal is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

usage in its two dimensions: quantitatively and qualitatively. It is the way combining both of corpus linguistics and educational linguistics. Among the aims of the study is to conciliate between educational books and the educational environment where the student finds himself forced to learn strange and rare vocabulary, structure and styles that he doesn't need, which affects his ability to practice and master language. Among the goals achieved by the research is developing a parallel corpus between Turkish and Arabic language, processing its material in order to teach Arabic to Turkish native speakers, providing a methodological vision of how to use this corpus in teaching Arabic as a second language and providing an easy way that helps to benefit from educational resources in saving time and effort spent in teaching and learning.

Keywords: Parallel Corpus; Educational Linguistics; Teaching Arabic for Non-Native Speakers; Turkish.

مُقدِّمة.

يسعى الباحثون في تعليم اللُّغات والمعنيُّونَ بطرائق تدريسها إلى ابتكار أساليبَ جديدة تُمكِّن المُتعلِّمين من تحصيل القدر الأكبر من المعرفة اللُّغويَّة المُرتبطة بواقع اللُّغة، وتُساعدُ المُعلِّمينَ في الوقت ذاته على توجيه طاقاتهم على نحو سليم. ومع تزايُد العناية بتعليم العربيَّة وتعلُّمها في المُجتمَع التُّركيّ، تسعى هذه الدِّراسة إلى بناء وتطوير مُدوَّنةٍ لُغويَّةٍ مُتوازية بينَ التُّركيَّة والعربيَّة، ثُمَّ توظيف مُخرَجات هذه المُدوَّنة في تعليم العربيَّة للنَّاطقينَ بالتُّركيَّة.

تَتَّبعُ هذه الدِّراسةُ منهجًا وصفيًّا يقومُ على جمع بياناتٍ لُغويَّة مُستمدَّة من واقع اللُّغَتَين (التُّركيَّة والعربيَّة)، ثُمَّ إخضاع هذه البيانات للمُعالجة الآليَّة، بهدف استخلاص مُعطياتٍ يُمكنُ توظيفُها في تعليم العربيَّة للنَّاطقينَ بغيرها. وتهدفُ الدِّراسةُ إلى:

إثراء المُحتوى العربيّ الرَّقميّ بموردٍ لُغويّ تعليميّ يتمثَّلُ في مُدوَّنةٍ لُغويَّةٍ مُتوازية بينَ التُّركيَّة والعربيَّة، واستثمار هذا المورد في تحقيق جدوى اقتصاديَّة تُساعدُ على توفير الوقت والجُهد المبذولَين في تعليم العربيَّة وتعلُّمها كلغةٍ ثانية، من خلال مُخرَجاتٍ لُغويَّةٍ أكثرَ فاعليَّة وقابليَّة للتَّطبيق.

توجيه القائمينَ على تعليم العربيَّة إلى ضوابطَ منهجيَّةٍ لإعداد الموارد التَّعليميَّة المُوجَّهة لمُتعلِّمي العربيَّة عبرَ المُستويات المُختلفة.

تقديم رُؤية لكيفيَّة الإفادة من المُدوَّنات المُتوازية في خلق بيئةٍ تفاعُليَّة للُّغة المُجتمع، ثُمَّ استثمار هذه البيئة في توليد نُصُوص اللُّغة الثَّانية الَّتي يستهدفُها المُتعلِّم.

وتسعى الدِّراسةُ إلى الإجابة عن الأسئلة الآتية:

ما مدى تأثير المُجتَمَع المُعَيَّن في اكتساب اللُّغة الثَّانية لأبناء هذا المُجتَمَع؟

كيفَ نبني مُدوَّنةً لُغويَّةً مُتوازيةً بينَ التُّركيَّة والعربيَّة؟ وكيفَ نستفيدُ من مُخرَجات هذه المُدوَّنة في تعليم العربيَّة للنَّاطقينَ بالتُّركيَّة وفقَ الإطار الأوروبيّ المرجعيّ العامّ للُّغات؟

كيفَ يُمكنُ توظيفُ لُغةٍ مُعيَّنةٍ في تعليم لُغاتٍ أخرى؟ وكيفَ يُمكنُ توظيفُ اللُّغات الأخرى (كالتُّركيَّة) في تعليم العربيَّة حالَ الحاجة إلى استخدامها في بيئةٍ غير عربيَّة؟

ما مُخرَجاتُ المُدوَّنة اللُّغويَّة المُتوازية بينَ التُّركيَّة والعربيَّة؟ وإلى أيّ مدًى يُمكنُ المُوائمة بينَ هذه المُخرَجات ومُستويات مُتعلِّمي العربيَّة و A, B, C؟

ما الإجراءاتُ الَّتي يُمكنُ اتِّخاذُها لأجل تطوير مُخرَجات المُدوَّنة اللُّغويَّة المُتوازية واستثمارها في تعليم العربيَّة للنَّاطقينَ بغيرها في كُلّ مُستوًى تعليميّ على حِدة؟

١. تمهيد نظريّ.

١.١. في ماهيَّة المُدوَّنات اللُّغويَّة المُتوازية.

أضحَت المُدوَّناتُ اللُّغويَّةُ Linguistic Corpora أحدَ أهمّ الموارد المُستخدَمة في ميادين: اللُّغة والتَّربية والمعلوماتيَّة، نظرًا لجدواها الاقتصاديَّة من ناحية، وتمثيلها لواقع المُجتَمَع اللُّغويّ من ناحيةٍ أخرى. وتُصنَّفُ المُدوَّناتُ اللُّغويَّةُ إلى عدَّة تصنيفاتٍ وفقًا لاعتباراتٍ مُختلفة، كاعتبار هيئة النُّصُوص؛ حيثُ تُصنَّفُ إلى المُدوَّنات النَّصِّيَّة والمُدوَّنات المنطوقة، واعتبار طبيعة النُّصُوص؛ حيثُ تُصنَّفُ إلى المُدوَّنات العامَّة والمُختَصَّة، واعتبار التَّنوُّع اللُّغويّ؛ حيثُ تُصنَّفُ إلى المُدوَّنات أحاديَّة اللُّغة والمُدوَّنات ثُنائيَّة اللُّغة ومُتعدِّدة اللُّغات[1]. وتُصنَّفُ المُدوَّناتُ اللُّغويَّةُ وفقَ كلِّ اعتبارٍ إلى أنواعٍ أخرى، كأن تُصنَّفَ المُدوَّناتُ مُتعدِّدة اللُّغات — مثلًا — إلى المُدوَّنات اللُّغويَّة غَير المُتوازية Non-parallel Corpora والمُدوَّنات اللُّغويَّة المُتوازية Parallel Corpora؛ وهذا النَّوعُ الأخيرُ هُو الَّذي ترتكزُ عليه الدِّراسة.

يُمكنُ تعريفُ المُدوَّنات المُتوازية بأنَّها كُتلة من النُّصُوص المُستمدَّة من واقع لُغةٍ طبيعيَّة وما يُقابلُها من نُصُوصٍ مُترجمة عنها في لُغةٍ أخرى. وتكونُ النُّصُوص أصلاً في إحدى اللُّغَتَين، وتُسمَّى (اللُّغة المصدر Source Language)، وترجمةً في اللُّغة الأخرى، وتُسمَّى (اللُّغة المُستهدَفة Target Language). وفي بعض الأحيان تشتملُ المُدوَّنةُ المُتوازية على نُصُوص أكثر من لُغَتَين، فتكونُ مصدرًا في إحدى اللُّغات وترجمةً في اللُّغات الأخرى[2]. ويُمكنُ التَّمثيلُ على المُدوَّنات المُتوازية بنصّ القُرآن الكريم وترجمة معانيه؛ حيثُ تكونُ العربيَّةُ مصدرًا لغيرها من اللُّغات، وكذلك العهد القديم وترجماته؛ حيثُ تكونُ العبريَّة والآراميَّة مصدرًا للُّغات الأخرى. ومن الأمثلة المُعاصرة المادَّة الصَّوتيَّة للأفلام السِّينمائيَّة حالَ ترجمتها.

ويُستفادُ من المُدوَّنات المُتوازية في ميادين البحث والصِّناعة والتَّدريس. فعلى مُستوى البحث اللُّغويّ والتَّربويّ، تُستخدمُ المُدوَّناتُ المُتوازية في الدِّراسات التَّقابُليَّة والمُقارَنة، وتُستخدَمُ أيضًا في علم التَّأثيل Etymology. وعلى مُستوى الصِّناعة، يُعتمَدُ على المُدوَّنات المُتوازية في صِناعة المعاجم ثُنائيَّة اللُّغة Bilingual Dictionaries وإنتاج أدوات التَّرجمة الآليَّة Machine Translation وإعداد المُقرَّرات التَّعليميَّة. أمَّا على مُستوى التَّدريس، فتُعَدُّ المُدوَّناتُ المُتوازيةُ موردًا لتطبيقات التَّعلُّم الآليّ Machine Learning، كما تُستخدَمُ على نطاقٍ واسع في تعليم اللُّغات الأجنبيَّة Teaching Foreign Languages، وغير ذلك من الميادين.

١.٢. توظيف اللُّغة المُعَيَّنة في تعليم لُغاتٍ أخرى.

يُعنى الباحثونَ في عُلوم اللُّغة والتَّربية بتوجيه مُتعلِّمي اللُّغة الثَّانية إلى مُعايشة هذه اللُّغة؛ ليسَ على مُستوى المهارات الأساسيَّة (القراءة، والكتابة، والتَّحدُّث، والاستماع) فحسب؛ وإنَّما أيضًا على مُستوى التَّفكير والاستنباط[3]. ويتَّفقُ هذا التَّوجيه مع نظريَّات اكتساب اللُّغة عُمُومًا. فالطِّفلُ — مثلًا — لا يتعلمُ لُغَتَهُ الأمّ عبرَ لُغةٍ ثانية، وإنَّما يكتسبُها من مُحيطِهِ المُمَثَّل في المُجتَمَع اللُّغويّ[4]؛ والحالُ كذلكَ لدى مُتعلِّمي اللُّغة الَّذينَ يعيشونَ في مُجتمعات هذه اللُّغة، حيثُ يكتسبونَها من خلال التَّمرُّس والمُعايَشة.

تبدو آثارُ هذه الرُّؤية حينَ ترتبطُ الغاية من تعلُّم اللُّغة بالرَّغبة في التَّواصُل مع أهلها والتَّعرُّف على ثقافاتهم[5]. لكنَّ هذا الأمرَ ليسَ مُطَّرِدًا. فأحيانًا نلجأ إلى تعلُّم لُغةٍ ما في غير بيئتها كلُغةٍ مُشتَرَكة بينَ جاليات كثيرة، وأحيانًا نتعلمُ اللُّغةَ لأغراضٍ بحْتيَّة محدودة أو لأغراضٍ دينيَّة. وللتَّمثيل على ذلك، يلجأ كثيرٌ من العرب المُقيمينَ في دُوَل الخليج العربيّ إلى تعلُّم اللُّغة الإنجليزيَّة، لأنَّها الأكثرُ انتشارًا بينَ أبناء الجاليات هناك. وغاية كثيرين منهم أن يتمكَّنوا من استخدام اللُّغة في الفهم والإفهام، لأنَّ الإنجليزيَّة لُغةٌ مُشتَرَكة بينَ العربيّ والصِّينيّ — مثلًا. وكلاهما ينشدُ استعمالَ لُغةٍ غير عربيَّة (هي الإنجليزيَّة) في مُجتَمَعٍ عربيّ.

حينَ يتعلمُ العربيُّ الإنجليزيَّة لمثل هذا الهدف، فإنَّه سيكونُ بحاجةٍ إلى مُعايشة الإنجليزيَّة في غير بيئتها. لكنَّهُ قد يجدُ نفسَهُ — في مراحل التَّعلُّم الأوَّليّ — أمامَ عباراتٍ تبدو غيرَ مُناسبة للبيئة العربيَّة الَّتي تَستخدمُ الإنجليزيَّة كلُغةٍ مُشتَرَكة؛ كتلك العبارات المُدرَجة في (الجدول ١).

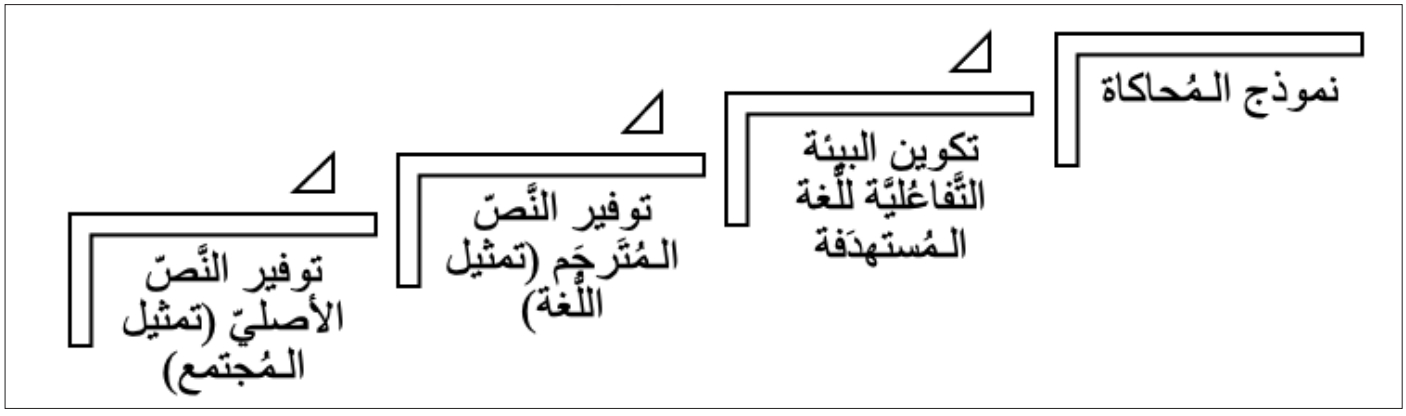
1 Danielle Barth- Stefan Schnell, *Understanding Corpus Linguistics* (England: Routledge, 2021), 200.
2 Marcos Zampieri - Preslav Nakov, *Similar Languages, Varieties, and Dialects: A Computational Perspective* (Cambridge, England: Cambridge University Press, 2021), 236.
٣ الخُولي، مُحمَّد عليّ. مدخل إلى علم اللُّغة، (عمَّان: دار الفلاح للنَّشر والتَّوزيع، ٢٠١٥)، ٢٠٤.
٤ القاسمي، عليّ. لُغة الطِّفل العربيّ «دراسات في السِّياسة اللُّغويَّة وعلم اللُّغة النَّفسيّ»، (بيروت: مكتبة لُبنان ناشرُون، ٢٠٠٩)، ٣٨.
5 Lynne Flowerdew, *Corpora and Language Education* (New York City: Springer, 2012), 132.

الجدول ١: نماذج لعباراتٍ – بينَ الإنجليزيَّة والعربيَّة – غير مُناسبة للبيئة العربيَّة

م	العبارة (بالإنجليزيَّة والعربيَّة)
1	I prefer red wine to white wine. أنا أُفضِّلُ النَّبيذَ الأحمرَ على النَّبيذ الأبيض.
2	Did your mother cooked a ham or a turkey for Christmas? هل قامَت والدتُكَ بطهي لحم الخنزير أم الدِّيك الرُّوميّ لاحتفال عيد الميلاد؟
3	If you carve a pumpkin for Halloween, you will be funny. لو قُمتَ بنحت اليقطين (القرع) لأجل الهالوين (عيد الرُّعب)، فسوفَ تبدو مُضحكًا.

وعلى مُستوى اللُّغة العربيَّة، فالأمرُ يبدو أكثرَ وُضوحًا؛ حيثُ أدَّت الاضطراباتُ الحادثة في المنطقة العربيَّة إلى هجرة كثيرين من العرب، فتنامَت بذلكَ أعدادُ الجاليات العربيَّة في بُلدانٍ غير عربيَّة، مثل: تركيا، وألمانيا، وكندا، بنسبٍ مُتفاوتة. وظهرَت هنا غايةٌ اجتماعيَّةٌ جديدة، ترتَّبت عليها الحاجة إلى تعلُّم العربيَّة للتَّواصُل بها مع العرب أنفسهم في مُجتمَع غير عربيّ. وبحُكم البيئة، فلن يكونَ من المُناسب أن يتعلمَ الأتراكُ العربيَّةَ – مثلًا – عبرَ مُفرداتٍ تنفردُ باستخدامها البيئةُ العربيَّة، لا سيَّما مُفرداتُ البيئة الصَّحراويَّة أو مُفردات اللُّغة العربيَّة في صُورَتِها: القديمة والوسيطة أو نحو ذلك. ولعلَّ مُطالعةً مُتأنِّيةً للمُقرَّر التَّعليميّ (العربيَّة بينَ يَدَيك) – المُوجَّه للنَّاطقينَ بغير العربيَّة – تجعلُنا نُدركُ حجمَ المُشكلة؛ حيثُ يجدُ المُتعلمُ غيرُ العربيّ نفسَه أمامَ موضُوعاتٍ تضُمُّ كلماتٍ وعباراتٍ تُعبِّرُ عن مُجتَمع الجزيرة العربيَّة تحديدًا[٦]، فيستشعر بذلك الفجوةَ القائمة بينَ اللُّغة والمُجتَمَع، ممَّا يُعيقُه عن اكتساب العربيَّة على الوجه المنشود. ولأجل مثل هذه الغايات، فقد يكونُ من المُفيد أن نخرُجَ عن الرُّؤية المثاليَّة لتعليم اللُّغة المُعيَّنة في إطار مُحيطها اللُّغويّ إلى رُؤيةٍ أخرى، نستثمرُ فيها اللُّغةَ الأمَّ للمُتعلِّم في بناء تصوُّرٍ للُّغة المُستهدَفة في غير مُحيطها. وبتوظيف ذلك على اللُّغة العربيَّة، يُمكنُ القولُ إنَّنا نسعى إلى بناء نموذج مُحاكاة للُّغة العربيَّة، حالَ استخدامها في مُجتمَع غير عربيّ، مُمثِّلينَ على ذلكَ بالمُجتَمَع التُّركيّ الَّذي تتنامى فيه الجالياتُ العربيَّة بصورةٍ كبيرة.

وحتَّى تتَّضِحَ الفكرةُ الَّتي تصبو إليها الدِّراسة، ينبغي التَّنويهُ إلى أنَّنا لا ننشُدُ توجيهَ المُتعلِّم لاستخدام لُغته (التُّركيَّة) في تعلُّم العربيَّة؛ وإنَّما ننشُدُ توظيفَ لُغة المُتعلِّم (التُّركيَّة) في بناء نموذج مُحاكاة، تخضعُ فيه اللُّغة العربيَّةُ لمُعطيات البيئة التُّركيَّة، لتبدو وكأنَّها أصيلة في هذه البيئة. ومن ثمَّ، يُمكنُ استخدامُ هذا النَّموذج في مُباشَرة عمليَّتَي التَّعليم والتَّعلُّم. ويُوضِّحُ (الشَّكل ١) تصوُّرًا للمراحل الَّتي يُمكنُ من خلالها خلقُ هذا النَّموذج.



الشَّكل ١: تصوُّر مراحل بناء نموذج مُحاكاة لتعليم اللُّغة الثَّانية (من الأسفل إلى الأعلى)

١.٣. المُدوَّنات المُتوازية وتعليم العربيَّة للنَّاطقينَ بغيرها.

حينَ نصنعُ مُدوَّنةً لُغويَّةً مُتوازية، فإنَّنا نُراعي طبيعةَ (اللُّغة المُستهدَفة) وأنظمتَها اللُّغويَّة: الصَّرفيَّة والتَّركيبيَّة والدِّلاليَّة؛ ونضطرُّ أحيانًا إلى مُراعاة جوانبَ غير لُغويَّة، كسياق الموقف أو السِّياق العاطفيّ، بهدف الوُصُول إلى أقرب المعاني المقصودة في نُصُوص (اللُّغة المصدر). ويعني هذا أنَّنا لا نقومُ بترجمةٍ حرفيَّة لنُصُوص اللُّغة المصدر؛ وإنَّما نقومُ بترجمةٍ أقربَ إلى المُحاكاة الَّتي تنتقلُ من خلالها اللُّغةُ والثَّقافةُ معًا[٧]. ولتوضيح ذلك نقول: إنَّنا نُترجمُ التَّعبيرَ الإنجليزيَّ (clouds in the She is) إلى العربيَّة ليكونَ على الصُّورة (إنَّها شاردة الذِّهن) خلافًا للتَّرجمة الحرفيَّة (إنَّها في الغُيُوم) الَّتي لا يُفهَمُ منها المقصود. ومع هذا، فإنَّ التَّصرُّفَ في التَّرجمة يكونُ محدودًا في هذا النِّطاق، نظرًا لتأثيره على الفهم والإفهام. أمَّا الجوانبُ الأخرى الَّتي تتَّصلُ بمباني اللُّغة وتراكيبها فتبدو أكثرَ نمطيَّة واتِّساقًا في اللُّغة المُستهدَفة؛ وهُو أمرٌ ضروريٌّ تفرضُهُ الحاجة إلى مُعالجة النُّصُوص آليًّا وإحصائيًّا، بهدف استخلاص الموارد التَّعليميَّة المُوَجَّهة للمُتعلِّم بصورةٍ مُباشرة، كالمُقرَّرات والمعاجم التَّعليميَّة.

وتبعًا لما ذُكِرَ آنفًا بشأن وجود غاياتٍ ناتجةٍ عن تعلُّم العربيَّة في غير بيئتها، كالأغراض الاجتماعيَّة والدِّينيَّة، فإنَّ ثمَّةَ سؤالًا يطرحُ نفسَه حولَ استخدام المُدوَّنات اللُّغويَّة المُتوازية في تعليم العربيَّة للنَّاطقينَ بغيرها (من الأتراك)؛ وهو (أيُّ اللُّغتَين تكونُ «المصدر»، وأيُّهما تكونُ «الهدف» ؟!). ستكونُ الإجابةُ البدهيَّة أنَّ لُغةَ المُتعلِّم (التُّركيَّة) هي المصدر، بينما تكونُ العربيَّة الهدف؛ وذلكَ لأنَّنا ننشُدُ محاكاةَ بيئة المُتعلِّم في لُغةٍ أخرى. ومع هذا، فإنَّ هذا الأمرَ لا يكونُ مُطِّردًا؛ إذ إنَّ غايةَ التَّعلُّم هي الَّتي تُحدِّدُ (المصدرَ والهدف). فإذا كانت الغايةُ هي التَّواصُلُ والمُعايَشة للعربيَّة في المُجتمَع التُّركيّ، فالتُّركيَّةُ هي المصدر؛ وإذا كانت الغايةُ دينيَّةً [للتَّعبُّد بالقُرآن – مثلًا – أو معرفة تفسيره أو فهم تعاليم السُّنَّة النَّبويَّة أو نحو ذلك] فالعربيَّةُ هي المصدر؛ والتُّركيَّةُ حينئذٍ هي الهدف. وإذا كانت الغايةُ عامَّةً بأن ينشُدَ المُتعلمُ تعلمَ العربيَّة لأغراضٍ مُتعدِّدة، فمن المُمكن أن نعتمدَ حينئذٍ على مُدوَّنةٍ لُغويَّةٍ تكونُ اللُّغةُ المصدرُ فيها لُغةً مُشتَرَكة (مثل الإنجليزيَّة) وتكونُ التُّركيَّةُ والعربيَّة مُتَرجَمَتَين عنها. وفي هذه الحالة، فإنَّنا لا نكونُ بحاجةٍ إلى النَّصّ الأصليّ، لأنَّهُ لا يُعبِّرُ عن لُغة المُتعلِّم، كما لا يُعبِّرُ عن اللُّغة الَّتي ينشُدُ تعلمَها. ومن ثمَّ، فإنَّ اللُّغةَ التُّركيَّة تقومُ مقامَ اللُّغة المصدر (وإن لم تكُن كذلك)، وتقومُ العربيَّةُ مقامَ اللُّغة الهدف.

١.٤. الدِّراسات السَّابقة.

وقَفَ الباحثُ على مجموعة من الدِّراسات الحديثة الَّتي تعرضُ لجوانب الإفادة من المُدوَّنات اللُّغويَّة عُمُومًا في تعليم اللُّغة الأجنبيَّة [الثَّانية]. وكانَ من أبرزها ما يأتي:

دراسة بعُنوان «التَّعلُّم بالاكتشاف في الفُصُول الدِّراسيَّة لترجمة اللُّغة: المُدوَّنات اللُّغويَّة كأدواتٍ مُساعدة للتَّعلُّم Discovery learning in the language-for-translation classroom: corpora as learning aids » للباحثة «سيلفيا برناردينيSilvia Bernardini». وقد نُشِرَت في عام ٢٠١٦م[٨].

كتاب «مناهج لسانيَّات المُدوَّنة في تعليم اللُّغة الثَّانية Foreign/Second Language Learning A Corpus-Linguistic Approach to » للباحث «ألفريد وارويك Warwick John Aflred ».

وقد نُشِرَ في عام ٢٠١٧م[٩].

٦ الفوزان، عبد الرَّحمن, حسين، مُختار, فضل، مُحمَّد. العربيَّة بينَ يَديك، سلسلة في تعليم اللُّغة العربيَّة لغير النَّاطقينَ بها، الطَّبعة الثَّالثة، (الرِّياض: العربيَّة للجميع، ٢٠٠٧)، ١/١٢٧.

7 Sue-Ann Harding - Ovidi Carbonell Cortés, *The Routledge Handbook of Translation and Culture* (England: Routledge, 2018), 121.

8 Scott Bernardini, "Discovery learning in the language-for-translation classroom: corpora as learning aids", *Cadernos de Tradução*, 36/1 (2016), 14-35.

9 Alfred Jones Warwick, *A Corpus-Linguistic Approach to Foreign/Second Language Learning: An Experimental Study of a New Pedagogicmodel for Integrating Linguistic Knowledge with Corpus Technology* (UK: Open Dissertation Press, 2017).

كتاب «المُمارَسة في تعلُّم اللُّغة الثَّانية Practice in Second Language Learning» للباحث «كريستيان جونز Christian Jones». وقد نُشِرَ في عام ٢٠١٨م [10].

كذلكَ فقد وقفَ الباحثُ على دراساتٍ تُعنى باستخدام المُدوَّنات في تعليم العربيَّة، منها:

دراسة بعُنوان «المُدوَّنة اللُّغويَّة لمُتعلِّمي العربيَّة ودورُها في تدريس العربيَّة للنَّاطقينَ بغيرِها Speakers in Teaching Arabic to Non-Native Learner Corpus and Its Potential Role Arabic » للباحثَين: «عبد الله الفيفي» وإيريك أتويل Eric Atwell». وقد نُشِرَت في عام ٢٠١٤م[11].

دراسة بعُنوان «المُدوَّنة اللُّغويَّة المُتوازية بينَ العربيَّة والإنجليزيَّة: مَورد جديد للتَّدريب على التَّرجمة وتعليم اللُّغة Training and Arabic-English Parallel Corpus: A New Resource for Translation Language Teaching» للباحثة هند العُتيبي. وقد نُشِرَت في ٢٠١٧م[12].

دراسة بعُنوان «التَّدريس القائم على المُدوَّنات اللُّغويَّة في الفُصُول الدِّراسيَّة العربيَّة Corpus-based teaching in the Arabic classroom» للباحثة «ميّ زكي». وقد نُشِرَت في عام ٢٠١٧م[13].

٢. إشكالات بناء المُدوَّنة المُتوازية لأغراض تعليم العربيَّة.

٢.١. نُدرة المُحتوى العربيّ الرَّقميّ.

لا يتجاوزُ المُحتوى العربيُّ الرَّقميّ Arabic Digital Content على الشَّبكة العنكبوتيَّة ٣ ٪ من إجمالي المُحتوى الرَّقميّ؛ علمًا بأنَّ ٢٠ ٪ فقط من المُحتوى العربيّ أصيل. أمَّا نسبة الـ ٨٠ ٪ الأخرى فهي مُحتوًى منسوخٌ ومُتكرِّر[14]. وهو ما يعني أنَّ نسبة المُحتوى العربيّ الرَّقميّ الأصيل أقل من (½ ٪) من نسبة المُحتوى الرَّقميّ العالميّ. وهي نسبةٌ قليلةٌ للغاية، مُقارنةً بعدد مُتحدِّثي العربيَّة الَّذي يتجاوز ٤٢٠ مليون مُتحدِّث، فضلًا عن كون العربيَّة إحدى اللُّغات السِّتّ المُعتَمَدَة في الجمعيَّة العامَّة للأمم المُتَّحدة UNGA. ويترتَّبُ على هذه النُّدرة صُعوبةُ الحُصُول على مُحتوى رقميّ للمُدوَّنات اللُّغويَّة العربيَّة، وأصعبُ من ذلك الحُصُول على مُدوَّناتٍ لُغويَّةٍ مُتوازية. ومن ثمَّ، فإنَّ بناءَ مُدوَّنةٍ لُغويَّةٍ مُتوازيةٍ بينَ العربيَّة واللُّغات الأخرى يستدعي عملًا جهيدًا لفرقٍ بحثيَّة كبيرة الحجم، تقومُ باختيار نُصُوص اللُّغة المصدر، ثمَّ تعملُ على تحريرها وتدقيقها وترجمة مُحتواها إلى اللُّغة المُستَهدَفة، ثمَّ إخضاع ذلكَ كلِّه للمُعالجة الآليَّة لاستخلاص مُخرَجات هذه المُدوَّنات. ويبدو في الأفق إشكالٌ آخر، مفادُهُ أنَّ المؤسَّسات أو الفرق البحثيَّة الَّتي يُمكنُ أن تعملَ على بناء المُدوَّنات المُتوازية لن تسمحَ — في الغالب — للباحثينَ الأفراد باستخدام هذه المُدوَّنات بصورةٍ مجَّانيَّة؛ حتَّى أنَّ بعضَ المؤسَّسات تُوجِّهُ طاقاتها لإنتاج المُدوَّنات اللُّغويَّة ضمنَ مواردَ لُغويَّة Language Resourc مُتعدِّدة لأهدافٍ استثماريَّة ربحيَّة، على النَّحو الَّذي نجِدُهُ مثلًا في بعض موارد مُؤسَّسة LDC التَّابعة لجامعة بنسلفانيا.

ويُمكنُ التَّغلُّبُ على هذا الإشكال — جُزئيًّا — بالبحث عن بدائلَ أخرى للمُحتوى الرَّقميّ. ومن وُجهة نظر الباحث، فإنَّ البديلَ الأكثرَ مُناسبةً لتحقيق أهداف الدِّراسة يتمثَّلُ في اللُّجوء إلى المُحتوى الحُرّ Free Content الَّذي ينتجُ عن المنصَّات الإلكترونيَّة المفتوحة، مثل مجموعة منصَّات «ويكي Wiki»، من خلال عملٍ تفاعُليٍّ بينَ مُستخدمي الشَّبكة العنكبوتيَّة. وما يُميِّزُ نُصُوصَ هذا المُحتوى أنَّها تخضعُ للملكيَّة العامَّة، على النَّحو الَّذي يسمحُ باستخدامها وتطويرها لأغراض البحث العلميّ. وتتنوَّعُ أشكالُ هذا المُحتوى لتشملَ: الموسوعات الحُرَّة، والتَّرجماتِ العامَّة للصُّوَتيَّات والمرئيَّات، وترجمات الوثائق العامَّة للمُنظَّمات الدَّوليَّة، وغيرِها[15]. ومع ذلكَ، فإنَّ هذا المُحتوى يظلُّ مُقيِّدًا، لأنَّهُ لا يُتيحُ المساحة اللَّازمة لاختيار النُّصُوص؛ كما أنَّهُ يستدعي مُراجعةً لُغويَّةً قبلَ الشُّرُوع في مُعالَجَتِه.

٢.٢. تباينُ النِّظام اللُّغويّ بينَ العربيَّة والتُّركيَّة.

تنتمي اللُّغة العربيَّة إلى فصيلة اللُّغات السَّاميَّة Semitic languages، وتتفرَّعُ هذه الفصيلة عن أسرةٍ لُغويَّةٍ أكبر، هي أسرة اللُّغات الأفروآسيويَّة Afroasiatic languages الَّتي تُعرَفُ أيضًا بأسرة اللُّغات الحاميَّة السَّاميَّة Hamito-Semitic؛ بينما تنتمي اللُّغةُ التُّركيَّة إلى فصيلة اللُّغات التُّركيَّة[16] [الأتراكيَّة] Turkic languages الَّتي تتفرَّعُ عن أسرة اللُّغات الألطيَّة Altaic languages . ولكلٍّ من الفصيلَتين (السَّاميَّة والتُّركيَّة) سماتٌ تُميِّزُ لُغاتها عن لُغات الفصيلة الأخرى، وتفرضُ تبايُنًا في النِّظام اللُّغويّ بمُستوياته المُختلفة.

وعلى سبيل المثال، في مُستوى البنية، يُشارُ في اللُّغة العربيَّة بالضَّمير (أنتَ) لخطاب المُفرَد المُذكَّر، ويُشارُ بالضَّمير (أنتِ) لخطاب المُفرَد المُؤنَّث. أمَّا في اللُّغة التُّركيَّة، فيُشارُ إلى المُفرَد المُذكَّر والمُفرَد المُؤنَّث بضميرٍ واحد، هو (Sen). والأمرُ كذلك في صِيغَتَي الضَّمير (أنتم) و (أنتُنَّ)؛ حيثُ يُشارُ إليهما بالضَّمير)Siz(. أمَّا على مُستوى التَّراكيب، فالجُملة الفعليَّة العربيَّة تأخُذُ النَّمطَ (الفعل، الفاعل، الكلمة المُتمِّمة)، فنقولُ (يقرأ الطالب الصَّحيفة)؛ بينما تأخذُ في التُّركيَّة النَّمط (الفاعل، الكلمة المُتمِّمة، الفعل) فنقولُ في ترجمة الجملة العربيَّة (Öğrenci gazete okuyor)، فيأتي الفاعلُ (Öğrenci) في أوَّل الجُملة، ويأتي الفعلُ (okuyor) في نهايتها.

النِّظامُ اللُّغويُّ مُتباينٌ إذن بينَ العربيَّة والتُّركيَّة. ويُمثِّلُ هذا الأمرُ إشكالًا كبيرًا عندَ بناء مُدوَّنةٍ لُغويَّةٍ مُتوازيةٍ بينَ اللُّغَتَين، لأنَّنا لا نصنعُ قاعدة بياناتٍ بالمُفردات أو الأساليب المُجرَّدة أو القواعد التَّركيبيَّة؛ وإنَّما نصنعُ مُحاكاةً مُوازيةً بينَ اللُّغَتَين؛ بحيثُ تصيرُ مُدوَّنةُ كلِّ لُغةٍ على حدة مُمثِّلةً فعليًّا لواقع هذه اللُّغة ومُعبِّرة بدقَّة عن طبيعة نظامها اللُّغويّ وخصائصه. وحالَ اللُّجوء إلى التَّرجمة الآليَّة لإحدى اللُّغَتَين، سعيًا إلى توليد اللُّغة الأخرى، فإنَّ المُخرَجات تستدعي تدخُّلًا بشريًّا ليسَ باليسير. وقد لا نشعرُ بهذا الإشكال إذا كانت المُدوَّنة المُتوازيةُ الَّتي نصنعُها بينَ (العربيَّة والعبريَّة) اللَّتين تنتميان معًا إلى فصيلة اللُّغات السَّاميَّة، أو بينَ (الألمانيَّة والهولنديَّة) اللَّتين تنتميان إلى فصيلة اللُّغات الجرمانيَّة؛ حيثُ يؤدِّي توافُقُ النِّظامِ اللُّغويّ في الفصيلة الواحدة إلى تيسير عمل الآلة. ولهذا، فإنَّ بناءَ المُدوَّنة المُتوازية يستدعي تدخُّلًا بشريًّا أو تدريبًا عميقًا لأدوات التَّرجمة الآليَّة عبرَ تقنيات التَّعلُّم الآليّ Machine Learning كي نتمكَّنَ من إنتاج النُّصُوص في اللُّغة المُستَهدَفة على الوجه الَّذي يسمحُ بتحقيق الغاية المنشودة.

٢.٣. حداثة استخدام المُدوَّنات اللُّغويَّة في ميادين البحث اللِّسانيّ التَّربويّ.

يُعنى علمُ اللُّغة التَّربويّ Educational Linguistics بالبحث في الجوانب التَّطبيقيَّة الَّتي تسمحُ باستثمار مناهج البحث اللُّغويّ في تعليم اللُّغات[17]. ويشهدُ هذا المَيدانُ تطوُّرًا ملحوظًا؛ حيثُ أمكنَ من خلاله ابتكارُ العديد من النَّظريَّات التَّربويَّة الَّتي تحملُ مُقارباتٍ اقتصاديَّة ترمي إلى توفير الوقت والجُهد المبذولَين في عمليَّتَي التَّعليم والتَّعلُّم. ومع ظُهور لسانيَّات المُدوَّنة Corpus Linguistics في مُنتصف القرن العشرين، وازدهار البحث فيها على مُستوى اللُّغة العربيَّة منذُ مطلع الألفيَّة الثَّالثة، أمكنَ استحداثُ وسائلَ جديدة لتمكين الآلة من مُعالجَة مجموعات النُّصُوص الكبيرة، بهدف استخلاص المُفردات والتَّراكيب والأساليب الأكثر دورانًا وتمثيلًا لواقع اللُّغة[18].

وعلى الرَّغم من ذلك، فإنَّ ثمَّةَ فجوةً واضحةً بينَ العامليْنَ في مَيدانَي: (علم اللُّغة التَّربويّ) و (لسانيَّات المُدوَّنة)؛ حيثُ يميلُ العاملونَ في المَيدان الأوَّل إلى البحث في مناهج البحث التَّربويّ المُجرَّدة، دونَ العناية الكافية بميادين البحث البينيّ الَّتي تُساعدُ في تطوير مُخرجات التَّعلُّم وبناء الموارد اللَّازمة للتَّعليم، لا سيَّما تلك الَّتي ترتبطُ بتطويع الآلة لأغراض البحث أو استخدام أساليب التَّحليل الإحصائيّ. ويميلُ العاملونَ في المَيدان الآخَر إلى التَّركيز على تطوير أساليب المُعالجَة الآليَّة أو الإحصائيَّة لمجموعات النُّصُوص، دونَ العناية بالمُوجِّهات النَّظريَّة الَّتي تتحقَّقُ من خلالها الأهدافُ البحثيَّة. وقد أدَّى وُجُودُ هذه الفجوة إلى إعاقة الاستثمار الأمثل لمُخرَجات مُعالجَة المُدوَّنات اللُّغويَّة في تعليم العربيَّة؛ سواءٌ لأبنائها أو للنَّاطقينَ بغيرِها.

والأمرُ ليسَ كذلكَ في لُغاتٍ أخرى، كالإنجليزيَّة والألمانيَّة والفرنسيَّة مثلًا؛ حيثُ أمكنَ توظيفُ المُدوَّنات اللُّغويَّة في تعليم اللُّغات الثَّلاث، عبرَ تطوير مناهج التَّعلُّم من ناحية، وبناء الموارد اللُّغويَّة التَّعليميَّة من ناحيةٍ أخرى. وبطبيعة الحال، فإنَّ وفرةَ الدِّراسات والموارد المعنيَّة بتوظيف المُدوَّنات في تعليم لُغةٍ ما تُسهِّلُ على الباحثينَ مُعالجَةَ جوانب القُصُور وابتكارَ أساليبَ جديدة لتعليم هذه اللُّغة. وربَّما تجدُرُ الإشارةُ إلى أنَّ الأمرَ يخضعُ في كثيرٍ من الأحيان للرُّؤية الاقتصاديَّة لدى القائمينَ على تعليم لُغةٍ ما؛ حيثُ تُمثِّلُ الجدوى الاقتصاديَّة من استخدام مواردَ لُغويَّة مُعيَّنة دافعًا إلى تطوير هذه الموارد وتوظيفها على الوجه الأمثل.

10 Christian Jones, *Practice in Second Language Learning* (Cambridge, England: Cambridge University Press), 2018.
11 Abdullah Alfaifi - Eric Atwell, "Arabic Learner Corpus and Its Potential Role in Teaching Arabic to Non-Native Speakers" *In proceedings of the 7th Biennial IVACS conference* (Newcastle: 2014).
12 Hind Alotaibi, "Arabic-English Parallel Corpus: A New Resource for Translation Training and Language Teaching", *Arab World English Journal (AWEJ)* 8/3 (2017), 319-337.
13 Mai Zaki, "Corpus-based teaching in the Arabic classroom: theoretical and practical perspectives" *International Journal of Applied Linguistics* 27/2 (2017), 514-541.
١٤ سبيتي، فيديل. المحتوى الرقمي العربي على الإنترنت قلب نابض وكيان ضئيل، (لندن: صحيفة إندبندنت، النُّسخة العربيَّة، ٢٠٢٠)؛ https://www.independentarabia.com/node/117381،(الوصول: ١٠ ديسمبر ٢٠٢١).
15 Allison Malinowski - Robert Maloy, *Wiki Works: Teaching Web Research and Digital Literacy in History and Humanities Classrooms* (Maryland: Rowman & Littlefield, 2017), 17.
16 Asya Pereltsvaig, *Languages of the World* (Cambridge: Cambridge University Press, 2020), 124.
17 Carolin Fuchs at al., *Language Education in Digital Spaces: Perspectives on Autonomy and Interaction* (Berlin: Springer Nature, 2021), II.
18 Fanny Meunier at al., *The Cambridge Handbook of Learner Corpus Research* (Cambridge: Cambridge University Press, 2015), 7.

٣. منهجيَّة بناء وتطوير مُدوَّنة لُغويَّة مُتوازية بينَ التُّركيَّة والعربيَّة.

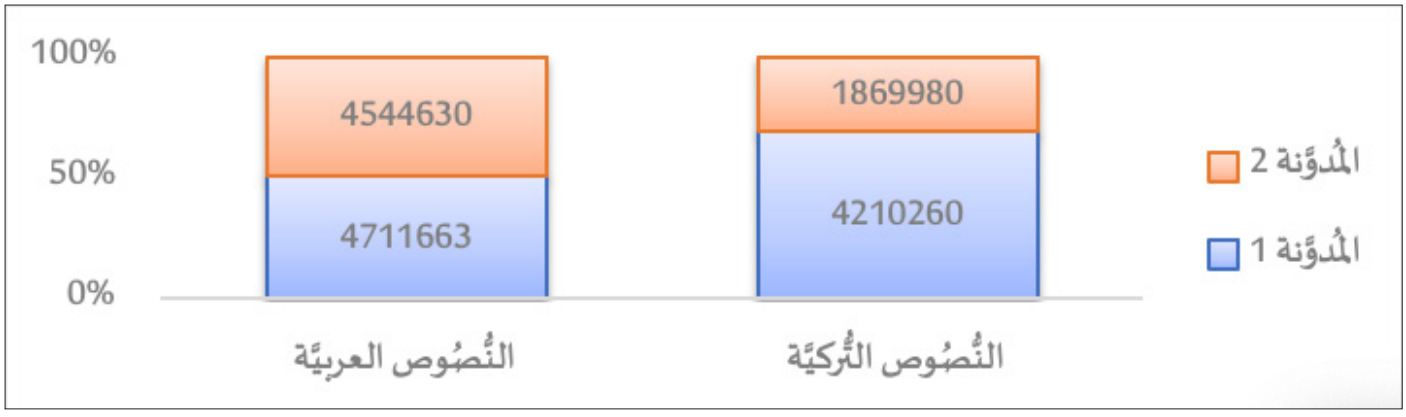
ننشُدُ من المُدوَّنة اللُّغويَّة المتوازية توفير مادَّة نصِّيَّة تُساعدُ على إنتاج الموارد التَّعليميَّة لمُتعلِّمي العربيَّة من أبناء المُجتَمَع التُّركيّ. ولأنَّ العربيَّة هي اللُّغة المُستَهدَفة، فالنُّصُوصُ التُّركيَّةُ ستكونُ مُنطلقًا مُثلًا لبيئة المُتعلِّم. ومن هذا المُنطلَق نستطيعُ استخلاصَ النُّصُوص العربيَّة الَّتي تُمثِّلُ واقعَ اللُّغة. وفي الطَّريق إلى إعداد منهجيَّةٍ واضحةٍ لبناء المُدوَّنة اللُّغويَّة المنشودة على النَّحو الَّذي تتحقَّقُ معهُ أهدافُ الدِّراسة، فقد وَضَعَ الباحثُ أربعةَ ضوابطَ لهذه المُدوَّنة؛ هي:

تنوُّع مادَّة المُدوَّنة بينَ المُعاصَرة والتُّراث.

تنوُّع مادَّة المُدوَّنة بينَ النُّصُوص العامَّة والمُختصَّة.

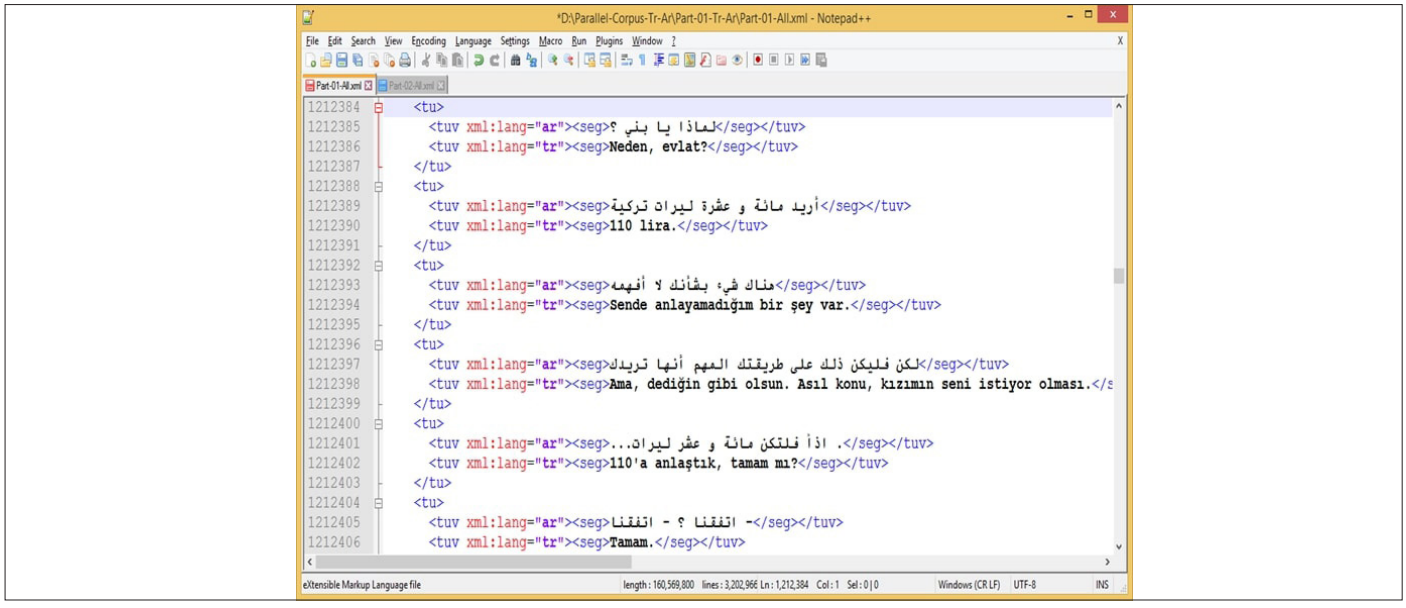
تنوُّع مادَّة المُدوَّنة بينَ النُّصُوص المنطوقة والنُّصُوص المكتوبة.

تنوُّع مادَّة المُدوَّنة بينَ النُّصُوص الحواريَّة [الجُمَل] والسَّرديَّة [الفِقرات].



الشَّكل ٢: مُخطَّط بيانيّ لنسبة النُّصُوص التُّركيَّة إلى العربيَّة في المُدوَّنة

ونظرًا للإشكالات الَّتي عَرَضَت لها الدِّراسةُ آنفًا، لا سيَّما ما يتعلَّقُ بنُدرة المُحتوى العربيّ الرَّقميّ، فقد استمدَّ الباحثُ مادَّةَ المُدوَّنة — موضوع الدِّراسة — من مصادرَ مُتعدِّدة، تَخضَعُ نُصُوصُها لرُخصة «جنو» للوثائق الحُرَّة GNU Free Documentation License الَّتي تسمحُ باستخدام وثائقها لأغراض البحث والتَّدريس[19]. وقسَّمَ المُدوَّنة إلى قسمَين، هما:

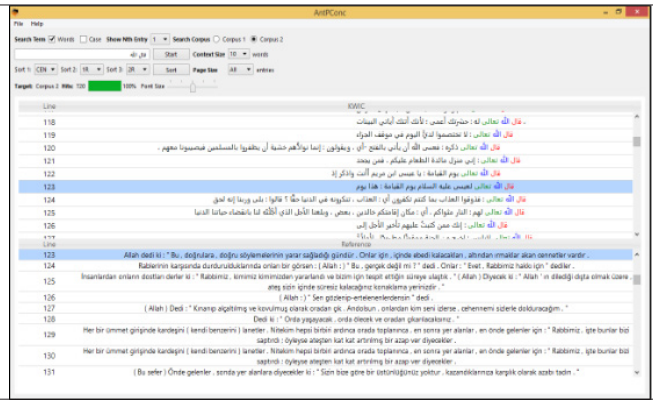


الشَّكل ٣: نموذج المُدوَّنة اللُّغويَّة على مِنَصَّة التَّحرير(Notepad++. v. 7.1)

المُدوَّنة ١: تُعبِّرُ عن اللُّغة المُعاصرة العامَّة، وتستمدُّ مادَّتَها من المادَّة الحواريَّة [المنطوقة] للمرئيَّات [تُركيَّة الأصل أو المنقولة إلى التُّركيَّة عن الإنجليزيَّة]، مع ترجماتها إلى العربيَّة. وهي مادَّةٌ صالحةٌ لتعليم العربيَّة للأغراض الاجتماعيَّة [للمُستوَيَين A, B].

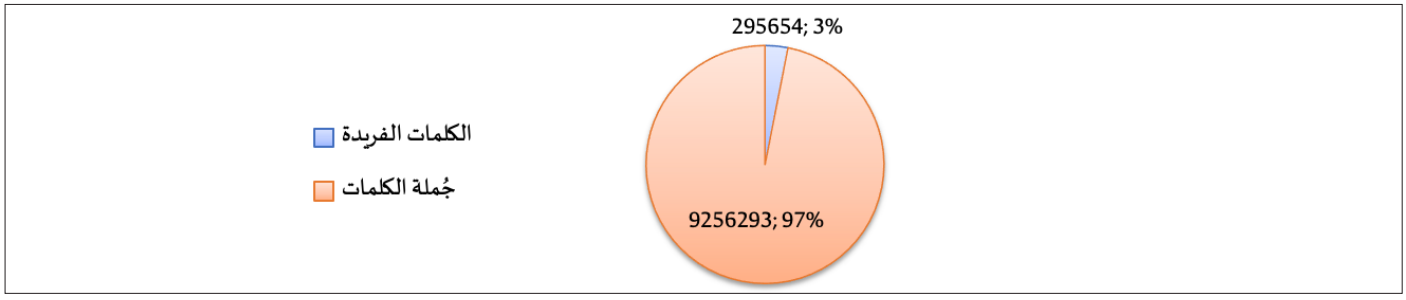
وقد بلغَ عددُ كلمات المُدوَّنة بقسميها (15336533) كلمة، مُوزَّعة على اللُّغتَين، بمجموع (6080240) كلمة تُركيَّة و (9256293) كلمة عربيَّة، بنسبة تقريبيَّة (2:3). وهي نتيجةٌ طبيعيَّة لاشتمال المُدوَّنة على قدرٍ من النُّصُوص التُّراثيَّة (المُدوَّنة ٢)؛ حيثُ تظهرُ مُرونةُ اللُّغة التُّركيَّة ومَيلها إلى الاختصار عندَ ترجمة هذه النُّصُوص من العربيَّة إلى التُّركيَّة.

19 The open parallel corpus (OPUS), "*OPUS multilingual corpora*" (Erişim 2 mayıs 2021).

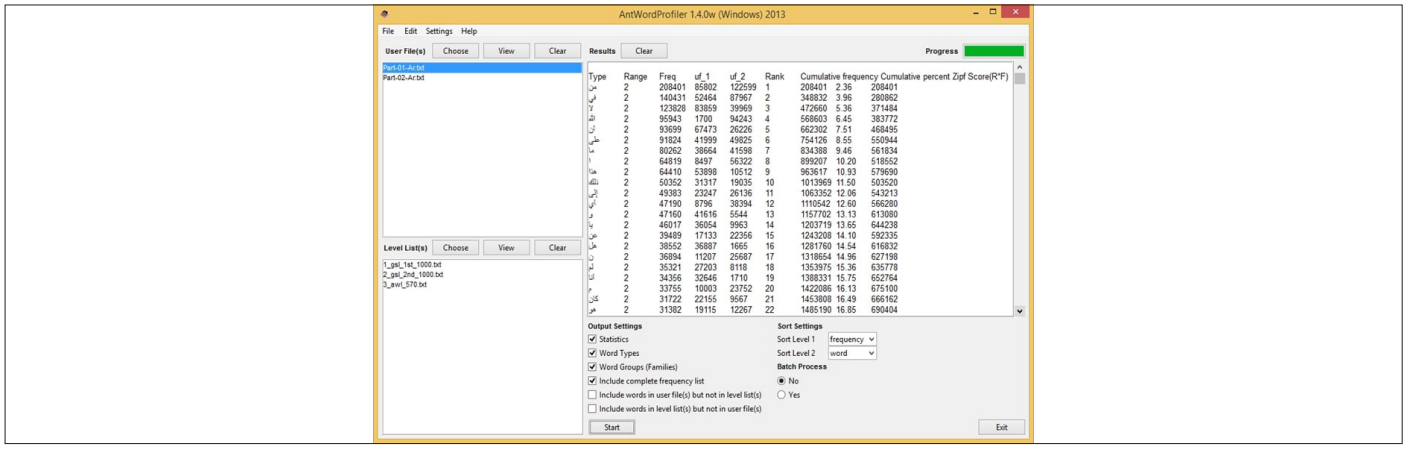


الشَّكل ٤: نموذج مُطابقة النُّصُوص بينَ التُّرْكِيَّة والعربيَّة باستخدام برمجيَّة (AntPConc)

ويُوضِّحُ (الشَّكل ٢) مُخطَّطًا بيانيًّا لنسبة النُّصُوص التُّرْكِيَّة إلى العربيَّة في المُدوَّنة اللُّغويَّة بقسمَيها: (المُدوَّنة ١) و (المُدوَّنة ٢)؛ حيثُ يُشيرُ إلى التَّقارُب النِّسبيّ بينَ عدد الكلمات التُّرْكِيَّة والعربيَّة في (المُدوَّنة ١) المُمثِّلة للُّغة المُعاصرة، وإلى التَّباين النِّسبيّ بينَ عدد الكلمات التُّرْكِيَّة والعربيَّة في (المُدوَّنة ٢) المُمثِّلة للُّغة التُّراثيَّة.



الشَّكل ٥: نسبة الكلمات الفريدة إلى جُملة الكلمات في المُدوَّنة العربيَّة



الشَّكل ٦: المُعالجَة الآليَّة للمُدوَّنة العربيَّة باستخدام برمجيَّة (AntWord Profiler)

كما يُوضِّحُ (الشَّكل ٣) نموذجَ المُدَوَّنة اللُّغويَّة في صُورتِها الخام Raw Corpus على مِنصَّة التَّحرير (Notepad++.v.7.1)[٢٠] ، قبلَ استخلاص النُّصُوص العربيَّة وإخضاعها للمُعالجَة.

الشَّكل ٣: نموذج المُدوَّنة اللُّغويَّة على مِنصَّة التَّحرير(Notepad++.v.7.1)

وقد قامَ الباحثُ بإخضاع المُدوَّنة للمُعالجة، بهدف مُطابقة الجُمل بينَ اللُّغتَين واستخلاص النُّصُوص العربيَّة، باستخدام برمجيَّة AntPConc الَّتي تُعنى بتحليل المُدوَّنات اللُّغويَّة المُتوازية إحصائيًّا؛ كما تُساعدُ على استخلاص نُصُوص كُلِّ لُغةٍ على حِدة، ثُمَّ مُوازاة هذه النُّصُوص بما يُقابلُها في اللُّغة الأخرى[٢١]. ويُوضِّحُ (الشَّكل ٤) نموذجًا لمادَّةٍ مُستَخلصة من المُدوَّنة اللُّغويَّة — موضوع الدِّراسة — بعدَ إخضاعها للمُعالجة الآليَّة.

وأعقَبَ الباحثُ ذلكَ باستخلاص النُّصُوص العربيَّة المُستهدَفة، ثُمَّ مُعالجتها إحصائيًّا لمعرفة نسبة الكلمات الفريدة إلى جُملة الكلمات العربيَّة، على النَّحو المُوضَّح في (الشَّكل ٥).

وسعيًا إلى تعيين المُفرَدات والتَّراكيب الَّتي ستكونُ مُوجِّهًا لمُعلِّمي العربيَّة ومُتعلِّميها، فقد قامَ الباحثُ بمُعالجَة النُّصُوص العربيَّة آليًّا؛ حيثُ أخضَعَها لبرمجيَّة AntWordProfiler الَّتي مَكَّنت من استخلاص قائمة المُفردات وتردُّداتها في المُدوَّنة[٢٢]، ثُمَّ أخضَعَها لتطبيق Nooj Concordance 3.2 المُتضَمَّن في مِنصَّة المُعالجَة الآليَّة NooJ v5.0، فأمكنَ بذلك استخلاصُ قائمة المُتلازمات والتَّراكيب[٢٣]، على النَّحو المُوضَّح في (الشَّكل ٦) و (الشَّكل ٧).

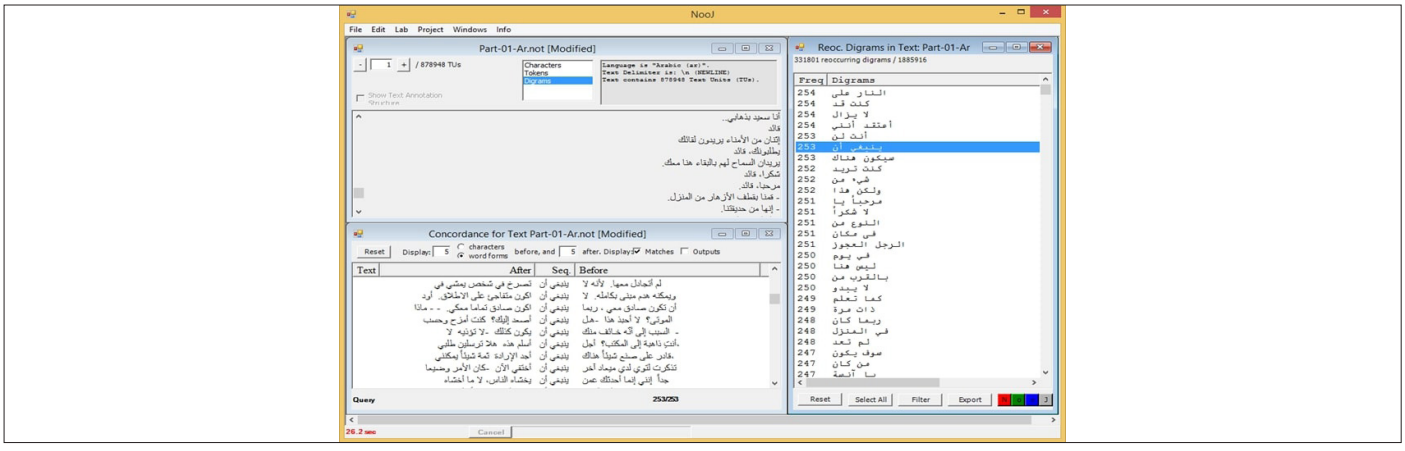
20 Don Ho, "*Notepad++: A free source code editor and Notepad replacement*" (Erişim 12 Aralık 2021).
21 Laurence Anthony, "*Antpconc: A freeware **parallel** corpus analysis toolkit for concordancing and text analysis*" (Erişim 15 Aralık 2021).
22 Laurence Anthony, "*AntWordProfiler: A freeware tool for profiling the vocabulary level and complexity of texts*" (Erişim 15 Aralık 2021).
23 Université de Franche-Comté, "*NooJ: A Corpus Processor - A Linguistic Development Environment - A Linguistic Engine for developing Natural Language Processing software Applications*" (Erişim 15 Aralık 2021).

٤. التَّطبيق.

حتَّى تتحقَّقَ الفائدةُ القُصوى من المُدوَّنة اللُّغويَّة، ينبغي توجيهُ مُخرَجات مُعالجتها لإعداد الموارد التَّعليميَّة المُباشرة، كالمُقرَّرات والمعاجم التَّعليميَّة، وفقَ إطارٍ قياسيٍّ يضمنُ الاستخدامَ الأمثلَ للنُّصُوص. ولأنَّنا ننشُدُ توظيفَ المُدوَّنة في تعليم العربيَّة للنَّاطقينَ بالتُّركيَّة، فإنَّ عملنا التَّطبيقيَّ يتمُّ وفقَ الضَّوابط المُعتَمَدة في الإطار الأوروبيّ المرجعيّ العامّ لُلغات Common European Framework of Reference for Languages (CEFR). ومن هُنا، يلزمُ التَّمييزُ بينَ المُستويات السِّتَّة (A1, A2, B1, B2, C1, C2)؛ ثُمَّ اختيار المُخرَجات المُناسبة لكلّ مُستوًى على حدة[٢٤] وحالَ التَّطبيق، يُمكنُ الجمعُ بيْن كُلِّ مُستويَيْن فرعيَّيْن؛ بحيثُ نصنعُ الموارِدَ التَّعليميَّة في ثلاثة أشكالٍ رئيسة، على النَّحو المُوضَّح في (الجدول ٢):

الجدول 2: توصيف مُخرَجات المُدوَّنة للمُستويات التَّعليميَّة A، B، C

م	المُستوى الرَّئيس	المُستويات الفرعيَّة	توصيف المُستوى	توصيف المُخرَجات		
				المُفرَدات	التَّراكيب	القواعد
1	A	A1	مُبتدئ	عامَّة أكثر شُيُوعًا	مألوفة قصيرة	أساسيَّة
		A2			يغلب عليها نُصُوص العبارات الحواريَّة يكونُ التَّركيزُ على احتياجات التَّواصُل الأساسيّ	
2	B	B1	مُتوسِّط	عامَّة شائعة	مألوفة قصيرة وطويلة	ثانويَّة
		B2			يغلب عليها نُصُوص العبارات الوصفيَّة يكونُ التَّركيزُ على احتياجات الفهم العامّ	
3	C	C1	مُتقدِّم	مُختصَّة أقلّ شُيُوعًا	مُتعدِّدة طويلة	حُرَّة
		C2			يغلب عليها نُصُوص المُتون والشُّرُوح يكونُ التَّركيزُ على احتياجات الفهم العميق	



الشَّكل ٧: المُعالجَة الآليَّة للمُدوَّنة العربيَّة باستخدام تطبيق (Nooj Concordance 3.2)

الجدول ٣: مُخرَجات المُدوَّنة وغاياتها للمُستويات التَّعليميَّة A، B، C

م	المُخرَج	المصدر	الغاية	المُستوى المُستهدَف		
				A	B	C
1	الأدوات الأكثر دورانًا	المُدوَّنة ١	تحديد أولويَّات القواعد. ربط الكلمات في الجُملة.	+		
2	الكلمات الأكثر دورانًا	المُدوَّنة ١ المُدوَّنة 2	بناء المُعجم التَّعليميّ. بناء المُقرَّر التَّعليميّ (A).	+	+	+
3	المُتلازمات الأكثر دورانًا	المُدوَّنة ١	فهم التَّراكيب والجُمَل. بناء المُقرَّر التَّعليميّ (B,A).	+	+	
4	النُّصُوص الحواريَّة	المُدوَّنة ١	تنمية المهارات الأساسيَّة. تنمية مهارات التَّواصُل. بناء المُقرَّر التَّعليميّ (A).	+		
5	النُّصُوص الوصفيَّة	المُدوَّنة ١	تنمية المهارات الثَّانويَّة. القُدرة على الفهم والإفهام. بناء المُقرَّر التَّعليميّ (C,B).		+	+
٦	نُصُوص المُتون والشُّرُوح	المُدوَّنة ٢	المُعايَشة الحُرَّة للُّغة. تنمية ذائقة التَّعبير. بناء المُقرَّر التَّعليميّ (C). مُمارَسة الشَّعائر الدِّينيَّة.			+

في ضوء هذا التَّوصيف، وبالنَّظر إلى طبيعة المُدوَّنة اللُّغويَّة المُتوازية بينَ التُّركيَّة والعربيَّة، يُمكنُ القولُ إنَّ بإمكاننا الإفادة من مُخرَجات (المُدوَّنة ١) في إعداد الموارد التَّعليميَّة للمُستوَيَيْن A, B؛ حيثُ تشتملُ على مادَّةٍ عامَّة من النُّصُوص الحواريَّة والوصفيَّة المُستخلَصة من المُسلسلات والأفلام السِّينمائيَّة والأفلام الوثائقيَّة. كما يُمكنُ الإفادة من مُخرَجات (المُدوَّنة ٢) في إعداد الموارد التَّعليميَّة للمُستوى C؛ حيثُ تشتملُ على مادَّةٍ مُختصَّة من النُّصُوص الدِّينيَّة. ويُوضِّحُ (الجدول ٣) بيانَ مُخرَجات المُدوَّنة وغايةَ كُلٍّ منها على حِدة.

وحتَّى نتبيَّنَ شكلَ هذه المُخرَجات على النَّحو الَّذي يُمكِّنُ من توظيفها في تعليم اللُّغة العربيَّة للنَّاطقينَ بغيرها (التُّركيَّة)، فإنَّنا سنعرضُ لمُخرَجات المُدوَّنة — تباعًا — في الجداول الآتية:

24 Rosamond Mitchell - Henry Tyne, *Language, Mobility and Study Abroad in the Contemporary European Context* (England: Routledge, 2021), 106.

الجدول 4: الأدوات الأكثر دورانًا في (المُدوَّنة 1)

م	الأداة / الكلمة الوظيفيّة	التَّردُّد في (المُدوَّنة 1)	النِّسبة المئويّة إلى جُملة كلمات (المُدوَّنة 1) [لأقرب 3 منازل عشريّة]
1	مِنْ / مَنْ	85755	0.783
2	لا	83744	0.771
3	في	80132	0.697
4	أنْ	66027	0.665
5	هذا	54031	0.626
6	على	41876	0.563
7	و	40745	0.545
8	ما	38776	0.511
9	هل	36914	0.495
10	يا	36334	0.489
11	أنا	32859	0.42
12	ذلك	31330	0.407
13	لقد	29484	0.783
14	لم	26535	0.771
15	هنا	25677	0.697
16	أنت	24072	0.665
17	إلى	23336	0.626
18	ماذا	23055	0.563
19	هناك	19773	0.545
20	نعم	19200	0.511

الجدول 5: الكلمات الأكثر دورانًا في جُملة الكلمات العربيّة [(المُدوَّنة 1) و (المُدوَّنة 2)]
[من غير الأدوات / من غير الكلمات الوظيفيّة]

م	الكلمة	التَّردُّد في جُملة الكلمات العربيّة	النِّسبة المئويّة إلى جُملة الكلمات العربيّة [لأقرب 3 منازل عشريّة]
1	الله	95714	1.034
2	كان	31962	0.345
3	كل	27984	0.302
4	شيء	19035	0.206
5	بعد	17045	0.184
6	كنت	16046	0.173
7	الآن	15181	0.164
8	يوم	14619	0.158
9	قال	14286	0.154
10	يجب	12310	0.133
11	قبل	11959	0.129
12	الدنيا	11630	0.126
13	الناس	11340	0.123
14	تعالى	11170	0.121
15	الأمر	10005	0.108
16	حسنًا	9781	0.106
17	الرسول	9561	0.103
18	الأرض	9279	0.1
19	بعض	9048	0.098
20	يكون	8516	0.092

الجدول 6: المُتلازمات الأكثر دورانًا في (المُدوَّنة 1)

م	المُتلازمة	التَّردُّد في (المُدوَّنة 1)
1	يجب أن	7347
2	أليس كذلك	4514
3	أنا لا	3784
4	من هنا	2975
5	كل شيء	2742
6	لم يكن	2695
7	ما الذي	2690
8	من قبل	2586
9	يمكن أن	2553
10	أريد أن	2442
11	من فضلك	2431
12	لا أعرف	2389
13	كل ما	2358
14	هل أنت	2293
15	هذا هو	2292
16	لا أستطيع	2273
17	من هذا	2195
18	أن يكون	1936
19	من أجل	1907
20	يبدو أن	1902

الجدول 7: نماذج من نُصُوص المُدوَّنة اللُّغويَّة العربيَّة

م	المُستوى	التَّوصيف	النَّمُوذج	المصدر
1	A	نُصُوص حواريَّة	كيف حال والدك؟ والدي مريض للغاية. فهمت. لديك أخٌ أكبر أيضًا؛ صحيح؟ نعم. كان في الحرب. هو الآن في المنزل. إذن فهو عاطل مثلي تمامًا. ألم تعد مدرسًا؟	المُدوَّنة 1
2	B	نُصُوص وصفيَّة	... قام (هتلر) بالزحف على (اليونان). ومرة أخرى بدون أي اعتراض ذي قيمة من جانب (موسكو). هذا الزحف لم يضعف من عزيمة (هتلر). بل ربما جعله أكثر إصرارًا على المواصلة، مع حصد جيوشه للمزيد من الانتصارات السهلة.	المُدوَّنة 1
3	C	مُثُون وشُرُوح	«يا بني إسرائيل اذكروا نعمتي التي أنعمت عليكم» بالشكر عليها بطاعتي « وأني فضَّلتكم » أي آباءكم «على العالمين» عالمي زمانكم. «واتقوا» خافوا «يوما لا تجزي» فيه «نفس عن نفس شيئًا» وهو يوم القيامة «ولا تقبل» بالتاء والياء «منها شفاعة» أي ليس لها شفاعة فتقبل (فما لنا من شافعين) «ولا يؤخذ منها عدل» فداء «ولا هم ينصرون» يمنعون من عذاب الله.	المُدوَّنة 2

٥. العمل المُستقبليّ.

لا يزعمُ الباحثُ أنَّ عملَهُ يخلُو من القُصُور؛ كما لا يزعمُ أنَّ المُدوَّنة اللُّغويَّة الَّتي صنَعَها تُعبِّرُ حقيقةً عن واقع المُجتمع التُّركيّ وطبيعة اللُّغة العربيَّة. ذلكَ أنَّ العملَ البحثيَّ الفرديَّ – مهما بلغَت دقَّةُ مُخرَجاته – لا يرقى إلى مُستوى العمل الجماعيّ المُنظَّم؛ لا سيَّما حينَ تكونُ الغايةُ إنتاجَ موردٍ تعليميّ يفي بحاجة فئاتٍ مُستهدَفةٍ كبيرةٍ نسبيًّا. أضف إلى ذلكَ أنَّ ابنَ اللُّغة [التُّركيَّة] هُو الأقدرُ على اختبار نُصُوص لُغته للتَّأكُّد من مُناسبة ترجمتها لمُستويات مُتعلِّمي اللُّغة المُستهدَفة [العربيَّة]. لهذا، فإنَّ الباحثَ ينشُدُ تطويرَ مُدوَّنته ومُخرَجاتها لتكونَ أكثرَ تعبيرًا عن حاجة الفئات المُستهدَفة؛ كما ينشُدُ توظيفَ هذه المُخرَجات في إنتاج مُقرَّرٍ تعليميٍ وافٍ، يلتزمُ بمعايير الإطار الأوروبيّ المرجعيّ العامّ للُّغات (CEFR) من ناحية، ويُوفِّرُ جُهُودَ المُعلِّمينَ من ناحيةٍ ثانية، ويُلبِّي طُمُوحَ المُتعلِّمينَ من ناحيةٍ ثالثة[٢٥]. وفي هذا الإطار، فإنَّ العملَ المُستقبليَّ للدِّراسة يستهدفُ مجموعةً من الإجراءات الَّتي نعرضُ لها في النِّقاط الآتية:

تنمية مُحتوى المُدوَّنة اللُّغويَّة لتشملَ ميادينَ أخرى، علميَّة وفنيَّة وأدبيَّة، تكونُ التُّركيَّةُ فيها أصلًا، وتكونُ العربيَّةُ ترجمةً عنها، سعيًا إلى تلبية احتياجات مُتعلِّمي العربيَّة من المُتحدِّثينَ بالتُّركيَّة في المُستوى المُتقدِّم (C) والرَّاغبينَ في تعلُّم العربيَّة لأغراضٍ خاصَّة.

تكوين فريق بحثيّ تفاعُليّ، يُشاركُ فيه عددٌ من الباحثينَ الأتراك المُجيدينَ للعربيَّة، بهدف تطوير المُدوَّنة اللُّغويَّة واختبار كفاءتها وتمثيلها للمُجتمَع التُّركيّ ومُراجعة ترجمتها إلى العربيَّة.

التَّواصُل مع المُؤسَّسات المعنيَّة بإنتاج الصَّوتيَّات والمرئيَّات النَّاطقة بالتُّركيَّة للفئات العُمريَّة المُختلفة، للمُساعدة في إثراء مُحتوى المُدوَّنة اللُّغويَّة بمادَّةٍ مكتوبةٍ ومنطوقة للأفلام والمُسلسلات والبرامج الحواريَّة وغيرها، وما يُقابلُها من ترجماتٍ عربيَّة.

الاستعانة بفريقٍ من الباحثينَ العرب في ضبط وتنسيق ومُراجعة مُخرَجات المُدوَّنة العربيَّة، بهدف تيسير إجراءات المُعالجَات الحاسوبيَّة والإحصائيَّة عليها، تمهيدًا للإفادة منها في بناء مواردَ تعليميَّةٍ مُتعدِّدة للمُستويات التَّعليميَّة A, B, C.

العملُ على استثمار مُخرَجات المُدوَّنة اللُّغويَّة لأجل إنتاج مواردَ تعليميَّةٍ ناجعة، تشملُ: مُقرَّراتٍ تعليميَّةٍ لمُتعلِّمي العربيَّة في مُستوياتها، ومُعجمًا تعليميًّا يُراعي دورانَ المُفرَدات في العربيَّة، ومادَّةً تعليميَّةً مسموعةً يُستفادُ منها في توجيه المُعلِّمينَ وتحديد أولويَّات المُتعلِّمين.

٦. نتائج الدِّراسة.

أبانَت الدِّراسةُ عن ماهيَّة المُدوَّنات اللُّغويَّة المُتوازية، وعرضَ الباحثُ لجوانب الإفادة منها في: (البحث، والصِّناعة، والتَّدريس).

كشَفَت الدِّراسةُ عن جدوى توظيف اللُّغة الأمّ للمُتعلِّم في تعليم اللُّغات الأخرى باستخدام المُدوَّنات المُتوازية، من خلال استثمار هذه المُدوَّنات في إيجاد نموذجٍ تفاعُليٍ لبيئة المُتعلِّم، ثُمَّ توليد اللُّغة المُستهدَفة عبرَ التَّرجمة. ومَثَّلَ الباحثُ على ذلك بنموذج المُجتمَع التُّركيّ حينَ يُسعى فيه إلى تعلُّم العربيَّة بُغيةَ استخدامها في هذا المُجتمَع.

عَرَضَت الدِّراسةُ لإشكالات بناء مُدوَّنة لُغويَّة مُتوازية لأغراض تعليم العربيَّة. وتمثَّلَت هذه الإشكالاتُ في: نُدرة المُحتوى العربيّ الرَّقميّ، وتباين النِّظام اللُّغويّ بينَ العربيَّة والتُّركيَّة، وحداثة استخدام المُدوَّنات اللُّغويَّة في ميادين البحث اللِّسانيّ التَّربويّ.

قدَّمَ الباحثُ منهجيَّةً لبناء وتطوير مُدوَّنة لُغويَّةٍ مُتوازية بينَ التُّركيَّة والعربيَّة بهدف توظيفها في تعليم العربيَّة للنَّاطقينَ بالتُّركيَّة. وقامَ الباحثُ ببناء هذه المُدوَّنة الَّتي اشتملَت على (١٥٣٣٦٥٣٣) كلمة، مُوزَّعة على اللُّغتَين، بمجموع (٦٠٨٠٢٤٠) كلمة تُركيَّة و (٩٢٥٦٢٩٣) كلمة عربيَّة، بنسبة تقريبيَّة (٢ : ٣).

على مُستوى التَّطبيق، أمكَنَ استخلاصُ النُّصُوص العربيَّة، وقامَ الباحثُ بإخضاع النُّصُوص العربيَّة للمُعالجَة الإحصائيَّة بهدف إعداد قائمةٍ بالكلمات الفريدة. وبلغَ عددُ هذه الكلمات (٢٩٥,٦٥٤) كلمة من جُملة الكلمات العربيَّة، بما نسبتُهُ ٣ ٪.

على مُستوى التَّطبيق، أمكَنَ إخضاعُ النُّصُوص العربيَّة للمُعالجَة الآليَّة. ونتجَت عن هذه المُعالجة مُخرَجاتٌ مُتعدِّدة، تشمل: الأدوات الأكثر دورانًا، والكلمات الأكثر دورانًا، والمُتلازمات الأكثر دورانًا، ونماذجَ النُّصُوص الحواريَّة، ونماذجَ النُّصُوص الوصفيَّة، ونماذجَ نُصُوص المتون والشُّرُوح. وقدَّمَ الباحثُ تصوُّرًا لكيفيَّة توظيف هذه المُخرَجات في تعليم العربيَّة للنَّاطقينَ بالتُّركيَّة في المُستويات A, B, C.

طرَحَت الدِّراسةُ رُؤيةً للعمل المُستقبليّ الَّذي يهدفُ إلى مُعالجَة القُصُور عبرَ إجراءاتٍ، منها: تنمية مُحتوى المُدوَّنة لتشملَ ميادينَ مُتعدِّدة، واختبار كفاءة المُدوَّنة للتَّأكُّد من تمثيل مادَّتها للمُجتمَع، واستثمار المُخرَجات في إنتاج مواردَ تعليميَّةٍ ناجعة.

الخُلاصة.

تبدو الفجوةُ واضحةً بينَ الموارد التَّعليميَّة المُوجَّهة لمُتعلِّمي العربيَّة النَّاطقينَ بغيرها والمُجتمَعات الَّتي يتعايشُونَ فيها مع العربيَّة حتَّى يتمكَّنوا من استخدامها [كلُغةٍ ثانية] والتَّواصُل بها على الوجه الأمثل. ويُمثِّلُ المُجتمَعُ التُّركيُّ نموذجًا على هذه الفجوة؛ حيثُ يفرضُ الواقعُ على مُتعلِّمي العربيَّة النَّاطقينَ بالتُّركيَّة أن يستخدموا العربيَّةَ في مُجتمعهم التُّركيّ الَّذي يشهدُ تنامي الجاليات العربيَّة. أضف إلى ذلك حاجة الأتراك إلى تعلُّم العربيَّة لأغراضٍ دينيَّة تتَّصلُ بمُمارسة الشَّعائر الإسلاميَّة؛ حيثُ تتجاوزُ نسبةُ المُسلمينَ في تُركيا ٩٦ ٪، وفقًا لإحصاءات الأرشيف الرَّقميّ (Wayback Machine). من هنا يرى الباحثُ ضرورةَ الانطلاق من مُجتمَع اللُّغة الأمّ للمُتعلِّم، لا من مُجتمَع اللُّغة الثَّانية الَّتي ينشُدُ تعلمَها. وفي مُحاولةٍ لإثبات قابليَّة هذه الرُّؤية للتَّطبيق، فإنَّ هذه الدِّراسة تسعى إلى بناء مُدوَّنةٍ لُغويَّةٍ مُتوازية بينَ التُّركيَّة والعربيَّة، بحيثُ تكونُ التُّركيَّةُ فيها تمثيلًا لواقع مُجتمَع المُتعلِّم، وتكونُ العربيَّةُ تمثيلًا لواقع اللُّغة المُستهدَفة. وقد قامَ الباحثُ ببناء مُدوَّنةٍ مُتوازيةٍ كاملةٍ، واستخلَصَ مُخرَجاتها، وقدَّمَ تصوُّرًا لكيفيَّة الإفادة من هذه المُخرَجات في تعليم العربيَّة للنَّاطقينَ بالتُّركيَّة في المُستويات التَّعليميَّة A, B, C.

مراجع الدِّراسة:

الخُولي، مُحمَّد علي. مدخل إلى علم اللُّغة، عمَّان: دار الفلاح للنَّشر والتَّوزيع، ٢٠١٥.

سبيتي، فيديل. المحتوى الرقمي العربي على الإنترنت قلب نابض وكيان ضئيل، لندن: صحيفة إندبندنت، النُّسخة العربيَّة، ٢٠٢٠، (تاريخ الوصول: ١٠ ديسمبر ٢٠٢١)؛ https://www.independentarabia.com/node/117381.

25 Eva Lindgren - Janet Enever, *Early Language Learning: Complexity and Mixed Methods* (Bristol: Multilingual Matters, 2017), 305-314.

الفوزان، عبد الرَّحمن, حسين، مُختار, فضل، مُحمَّد. العربيَّة بينَ يَديك، سلسلة في تعليم اللُّغة العربيَّة لغير النَّاطقينَ بها، الطَّبعة الثَّالثة، الرِّياض: العربيَّة للجميع، ٢٠٠٧.

القاسمي، عليّ. لُغة الطِّفل العربيّ «دراسات في السِّياسة اللُّغويَّة وعلم اللُّغة النَّفسيّ»، بيروت: مكتبة لُبنان ناشرُون، ٢٠٠٩.

Hakem Değerlendirmesi: Dış bağımsız.

Çıkar Çatışması: Yazar çıkar çatışması bildirmemiştir.

Finansal Destek: Yazar bu çalışma için finansal destek almadığını beyan etmiştir.

Peer-review: Externally peer-reviewed.

Declaration of Interests: The author have no conflicts of interest to declare.

Funding: The author declared that this study has received no financial support.

KAYNAKÇA

Alfaifi, Abdullah - Atwell, Eric. "Arabic Learner Corpus and Its Potential Role in Teaching Arabic to Non-Native Speakers". *In proccedings of the 7th Biennial IVACS conference*. Newcastle: 2014. [Crossref]

Alfwzan, Abdul-Rahman at al. *el-'Arabiyye beyn yedeyk silsile fi ta'limi'l-luga'l-'Arabiyye li gayri natikine bihâ*. Riyad: el-'Arabiyyetu li'l-Cemî, 2007.

Alkhuwly, Muhammed Aly. *el-Medhal ila 'ilmi'l-luga*. Uman: Dâru'l-Felah li'n-Neşri ve't-Tevzî', 2015.

Alotaibi, Hind. "Arabic-English Parallel Corpus: A New Resource for Translation Training and Language Teaching". *Arab World English Journal (AWEJ),* 8/3 (2017), 319-337. [Crossref]

Alqasmy, Aly. *Luga tifli'l-Arabî dirâsât fi siyâseti'l-lugaviyye ve ilmi'l-luga'n-nefsî*. Beyrut: Mektebutu Lübnan Nâşirûn, 2009.

Anthony, Laurence. "Antpconc: A freeware **parallel** corpus analysis toolkit for concordancing and text analysis". Erişim 15 Aralık 2021. *http://www.laurenceanthony.net/software/antpconc.*

Anthony, Laurence. "AntWordProfiler: A freeware tool for profiling the vocabulary level and complexity of texts". Erişim 15 Aralık 2021. *http://www.laurenceanthony.net/software/antwordprofiler.*

Barth, Danielle - Schnell, Stefan. *Understanding Corpus Linguistics*. England: Routledge, 2021.

Bernardini, Scott. "Discovery learning in the language-for-translation classroom: corpora as learning aids". *Cadernos de Tradução* 36/1 (2016), 14-35. [Crossref]

Flowerdew, Lynne. *Corpora and Language Education*. New York City: Springer, 2012.

Fuchs, Carolin at al. *Language Education in Digital Spaces: Perspectives on Autonomy and Interaction*. Berlin: Springer Nature, 2021.

Harding, Sue-Ann - Cortés, Ovidi Carbonell. *The Routledge Handbook of Translation and Culture*. England: Routledge, 2018.

Ho, Don. "Notepad++: A free source code editor and Notepad replacement". Erişim 12 Aralık 2021. *https://notepad-plus-plus.org*.

Jones, Christian. *Practice in Second Language Learning*. Cambridge: Cambridge University Press, 2018.

Lindgren, Eva - Enever, Janet. *Early Language Learning: Complexity and Mixed Methods*. Bristol: Multilingual Matters, 2017.

Malinowski, Allison - Maloy, Robert. Wiki Works: *Teaching Web Research and Digital Literacy in History and Humanities Classrooms*. Maryland: Rowman & Littlefield, 2017.

Meunier, Fanny at al. *The Cambridge Handbook of Learner Corpus Research*. Cambridge: Cambridge University Press, 2015.

Mitchell, Rosamond - Tyne, Henry. *Language, Mobility and Study Abroad in the Contemporary European Context*. England: Routledge, 2021.

OPUS, The open parallel corpus. "OPUS multilingual corpora". Erişim 2 mayıs 2021. *http://opus.nlpl.eu*.

Pereltsvaig, Asya. *Languages of the World*. Cambridge: Cambridge University Press, 2020.

Sobaity, Fidel. "*el-Muhteva'r-rakmî al-Arabî 'ale'l-internet kalb nabiz ve kayyân*". London: Independent [Arabic version], 2020, Erişim 10 Aralık 2021. https://www.independentarabia.com/node/117381.

Université de Franche-Comté. "NooJ: A Corpus Processor - A Linguistic Development Environment - A Linguistic Engine for developing Natural Language Processing software Applications". Erişim 15 Aralık 2021. *http://www.nooj-association.org*.

Warwick, Alfred Jones. *A Corpus-Linguistic Approach to Foreign/Second Language Learning: An Experimental Study of a New Pedagogicmodel for Integrating Linguistic Knowledge with Corpus Technology*. UK: Open Dissertation Press, 2017: [Crossref]

Zaki, Mai. "Corpus-based teaching in the Arabic classroom: theoretical and practical perspectives". *International Journal of Applied Linguistics* 27/2 (2017), 514-541. [Crossref]

Zampieri, Marcos - Nakov, Preslav. *Similar Languages, Varieties, and Dialects: A Computational Perspective*. Cambridge: Cambridge University Press, 2021.

EXTENDED SUMMARY

This study follows a descriptive approach based on collecting linguistic data derived from the reality of the two languages (Turkish and Arabic), and then subjecting these data to computational processing, with the aim of extracting data that can be used in teaching Arabic to non-Arabic speakers. The study aims to enrich the digital Arabic content with an educational language resource represented in a parallel corpus between Turkish and Arabic, and to use this resource in achieving economic feasibility that helps save time and effort in learning Arabic and teaching it as a second language. It also aims to direct those in charge of teaching Arabic to methodological controls for preparing educational resources for learners of Arabic across the different levels. We seek from the parallel corpus to provide a textual material that helps produce educational resources for Arabic learners from the Turkish community. Because Arabic is the target language, Turkish texts will be a representative starting point for the learner's environment. From this point of view, we can derive the Arabic texts that represent the reality of the language. On the way to preparing a clear methodology for building the desired corpus in a way that achieves the objectives of the study, the researcher has set four controls for this corpus: Diversity the corpus texts between contemporary and heritage, combining between general and specialized texts, Diversity the corpus texts between the spoken texts and the written texts and Diversity the corpus between dialogic texts [sentences] and narration [paragraphs]. The number of words in the corpus in its two sections is (15,336,533) words, distributed in the two languages, with a total of (6,080,240) Turkish words and (9,256,293) Arabic words, with an approximate ratio of (2:3). It is a natural result of the corpus's inclusion of a certain amount of heritage texts. The flexibility of the Turkish language and its tendency to be short appear when translating these texts from Arabic into Turkish. The researcher processed the corpus using computational and statistical tools, with the aim of matching sentences between the two languages and extracting Arabic texts; then the researcher extracted the target Arabic texts, and then processed it statistically to find out the ratio of unique words (Words without frequency) to the sentence of Arabic words. The study revealed the feasibility of employing the learner's mother tongue (First language) in teaching other languages using parallel corpora, by investing these corpora in creating an interactive model for the learner's environment, then generating the target language through translation. The researcher exemplified this by the model of Turkish society when it seeks to learn Arabic in order to use it in this society. The study presented the problems of constructing a parallel corpus for the purposes of teaching Arabic. These problems were represented in scarcity of digital Arabic content, disparity of the linguistic system between Arabic and Turkish and the recent use of linguistic corpora in the fields of educational linguistic research. At the application level, it was possible to extract the Arabic texts, and the researcher subjected the Arabic texts to statistical processing in order to prepare a list of unique words. The number of these words reached (295,654) words from among the Arabic words, with a percentage of 3%. At the application level also, it was possible to subject Arabic texts to computational processing. This processing resulted in multiple outputs, including: the most revolving particles, the most revolving words, the most revolving collocations, models of dialogic texts, models of descriptive texts and models of texts and annotations. The researcher presented a vision of how to use these outcomes in teaching Arabic to Turkish speakers at levels A, B and C. The study presented a vision for future work that aims to solve the deficiencies through measures, including: developing the corpus to include multiple fields, testing the corpus efficiency to ensure that its material represents society and investing outputs in producing useful educational resources.