

תחכום לקסיקלי: שיטה להוראה מדורגת של אוצר המילים בעברית כשפה שנייה לערבית ולהערכתו

| איהאב ח'ליל אבורביעה*

מילות מפתח: רכישת שפה שנייה, אוצר מילים, הוראת עברית, רשימת מילים שכיחות, שיפוט מורים

קיים שיעור זהה של גיוון לקסיקלי (100%), שכן בשניהם מופיעות שלוש מילים ייחודיות מתוך שלוש מילים בסך הכול. עם זאת, קיים ביניהם פער מהותי: המשפט הראשון מכיל מילים בסיסיות השכיחות בשלבי למידה מוקדמים, ואילו המשפט השני כולל אוצר מילים מתקדם שאינו שכיח המעיד על מיומנות לקסיקלית גבוהה ביחס למשפט הראשון. הבדל זה מדגים את תרומתו הייחודית של מדד התחכום הלקסיקלי להערכת מיומנות לקסיקלית.

הגדרת תחכום לקסיקלי ופירוט שיטות הערכתו

תחכום לקסיקלי מוגדר כאחוז המילים המתוחכמות או המתקדמות בטקסט (Lindqvist et al., 2011). מדד זה משמש לכימות השימוש בפריטי אוצר מילים ששכיחותם נמוכה (Malvern et al., 2004), וכן לזיהוי שימוש במונחים טכניים ובז'רגון מקצועי המאפשרים הבעה מדויקת בתחום דעת ספציפי (Read, 2000). עם זאת, אין תמימות דעים באשר להגדרת המדויקת של מילים מתוחכמות או מתקדמות, מה שהוביל להתפתחות שתי שיטות הערכה עיקריות: האחת מתבססת על רשימות שכיחות והשנייה על שיפוט מורים.

בשיטה הכמותית המבוססת על רשימות שכיחות (שבהן מוצגות מילים לפי סדר שכיחותן בשפה – מהנפוצות ביותר לנדירות ביותר), מילים מתוחכמות מוגדרות כמילים נדירות בשפה (Vermeer, 2000; Kyle & Crossley, 2016), ובפרט כאלו הנמצאות מעבר ל-2000 המילים השכיחות ביותר. גישה זו נשענת על ממצאים שלפיהם רכישת אוצר מילים קשורה קשר הדוק לשכיחות המילים (Nation, 2001; Ellis, 2002; Tekmen & Daloğlu, 2006), ומילים שכיחות נרכשות בקלות רבה יותר ממילים נדירות (Tidball & Treffers-Daller, 2008). יתרה מכך, שימוש במילים נדירות משקף לא רק אוצר מילים עשיר ומיומנות לשונית גבוהה, אלא הוא גם אינדיקטור להצלחה אקדמית (Linnarud, 1986; Laufer & Nation, 1995).

ערך זה דן במדד התחכום הלקסיקלי ככלי להוראת אוצר מילים ולהערכתו תוך השוואה למדד לקסיקלי נוסף – גיוון לקסיקלי. הוא סוקר את השיטות הקיימות להערכת תחכום לקסיקלי – הן הגישה הכמותית המבוססת על שכיחות מילים בשפה הן הגישה האיכותנית המתבססת על שיפוט מורים – ומציג את מגבלותיהן. כמו כן, הערך מתמקד בהמלצות יישומיות להוראת אוצר מילים מדורגת ולהערכת הלקסיקון בעברית כשפה שנייה (L2) לדוברי ערבית בהתבסס על מדד זה.

הערכת ידע אוצר המילים של לומדי L2 מתבצעת לרוב בשתי דרכים עיקריות: מדידת גודל אוצר המילים (רוחב הידע) ומדידת איכות אוצר המילים (עומק הידע) (Ellis, 2015). מדד הגיוון הלקסיקלי מתמקד ברוחב הידע ומוודד את שיעור המילים השונות במדגם נתון (אבורביעה, 2020; Read, 2020; Jarvis, 2017; Abu-Rabiah, 2020, 2023), ואילו מדד התחכום הלקסיקלי (lexical sophistication) בוחן את עומק הידע של הלומד תוך ייחוס חשיבות שונה למילים בהתאם לרמת התקדמותן (Lindqvist et al., 2011).

שני המדדים האלה הם כלים חיוניים להערכת מיומנות לקסיקלית. עם זאת, גיוון לקסיקלי רב לכשעצמו אינו מעיד בהכרח על מיומנות לקסיקלית מתקדמת, שכן ניתן להשיגה באמצעות שימוש במילים פשוטות ושכיחות. לפיכך, מדד התחכום הלקסיקלי מספק הערכה מעמיקה יותר של מיומנותיו הלקסיקליות של הלומד, בהיותו מתבסס על אבחנה איכותית בין מילים ולא על ספירה כמותית שלהן (Daller et al., 2003). לדוגמה, בשני המשפטים האלה:

1. "סטודנט פתר שאלה".

2. "בוזז ניתץ זגוגית".

* ד"ר איהאב ח'ליל אבורביעה הוא ראש ההתמחות בלשון העברית כשפה שנייה במכללה האקדמית לחינוך ע"ש קיי. מחקריו עוסקים ברכישת שפה שנייה ובהערכתה, במיוחד אוצר מילים. אבורביעה מכהן כחבר בוועדת המקצוע עברית לבתי ספר ערביים במשרד החינוך. בתשפ"ה הוא נבחר כחבר בפורום הצעירים במדעי הרוח והחברה של האקדמיה הישראלית הלאומית למדעים. אבורביעה היה חוקר בתר-דוקטורט אורח, במחלקה לבלשנות, באוניברסיטת יוטה, ארה"ב, בסיוע מלגת פולברייט לחוקרי בתר-דוקטורט ומלגת רוטשילד של קרן יד הנדיב. <https://orcid.org/0000-0002-8837-1089>

יכולות לנוע מבסיסיות למתקדמות, תופעה הנכונה הן למילים שכיחות (כמו 'ילד', 'ילדים', 'ילדי') והן למילים נפוצות פחות (כמו 'לשכה', 'לשכות') (אבורביעה, 2022). שנית, תהליך הדירוג עצמו דורש משאבים ניכרים מצד המדרגים, שכן דירוג מאות תבניות מילים לפי סקאלה מרובת רמות הוא משימה מורכבת ותובענית (אבורביעה, 2022). מגבלה זו עלולה לפגוע בישימות השיטה בשדה החינוכי.

שיטה מותאמת יותר להערכת תחכום לקסיקלי

בשל המגבלות בשיטות הקיימות אבורביעה (Abu-Rabiah, 2024) מציע שיטה משופרת להערכת תחכום לקסיקלי המבוססת על שיפוט מורים למילים מחוץ להקשר. לפי שיטתו הדירוג מתבצע ברמת המילה הבודדת (כולל נטיותיה השונות), ולא רק ברמת התבנית או צורת היסוד הבסיסית, כשהמילים הן מחוץ להקשר. השיטה מציעה מסגרת דירוג היררכית בת ארבע רמות: (1) מילים שכיחות/בסיסיות ביותר, (2) מילים שכיחות/בסיסיות, (3) מילים נדירות/מתקדמות, (4) מילים נדירות/מתקדמות ביותר. הדירוג צריך להיות מותאם ספציפית ליכולות הצפויות מקבוצת לומדים מסוימת, תוך התחשבות בהשפעת L1. לדוגמה, המילה 'רומח' מסווגת כבסיסית עבור דוברי ערבית (בשל הקוגנט 'رُمح' - rumḥ) אך כמתקדמת עבור דוברי אנגלית.

הממצאים במחקר המציג שיטה חדשה זו הראו מתאם הפוך בין רמת התחכום הלקסיקלי לדיוק התרגום – ככל שהמילים היו מתוחכמות יותר מבחינה לקסיקלית ('משפחה' > 'טיפול' > 'התחייבות' > 'נחלה'), כך פחת דיוק התרגום (מ-L2 ל-L1) של דוברי ערבית הלומדים עברית. גישה זו מציעה כלי רגיש יותר להערכת אוצר מילים, המתייחס הן למאפיינים המורפולוגיים של L2 ולקרבנות ל-L1, הן לרמת השליטה המצופה מהלומדים.

יישום עקרונות התחכום הלקסיקלי בהוראה מדורגת של אוצר מילים ובהערכתו

על אף האתגרים הכרוכים בהערכתו, התחכום הלקסיקלי מציע כלי פדגוגי משמעותי לפיתוח הוראת אוצר מילים בעברית כ-L2 לדוברי ערבית ובהערכתו. יישום עקרונותיו יכול להתבצע במספר מישורים.

במישור פיתוח חומרי הלימוד, מומלץ לגבש רשימות מילים מדורגות המבוססות על שילוב בין נתוני שכיחות מהעברית הכתובה והדבורה (להרחבה ראו Abu-Rabiah, 2025) לבין שיפוט מורים – מורי עברית מנוסים. תהליך זה יאפשר יצירת מדרג לקסיקלי התואם את שלבי הרכישה השונים – ממילים בסיסיות בשכבות הגיל הצעירות, דרך מילים בינוניות בחטיבות הביניים, ועד לאוצר מילים אקדמי ומקצועי בחטיבה העליונה. חשוב לציין כי הוראת המילים

עם זאת, מחקרים עדכניים מצביעים על מגבלותיה של הגישה המבוססת אך ורק על שכיחות. נמצא כי תחכום לקסיקלי מושפע מגורמים נוספים מלבד שכיחות (Horst & Collins, 2006; Tidball & Treffers-Daller, 2008). כך למשל, מילים ב-L2 הדומות (ולפעמים אף זהות) בצליל ובמשמעות לשפת האם (L1) של הלומד (קוגנטים כגון 'תִּבְּן' ו-'tīb-n', להרחבה ראו אבורביעה, 2023) גם אם הן נדירות, נרכשות בקלות יחסית למילים נדירות אחרות שאינן דומות ל-L1 (Tidball & Treffers-Daller, 2008). לומדי L2 אף מעדיפים שימוש במילים כאלו גם אם הן נדירות על פני מילים שכיחות אך אינן דומות ל-L1 (Cobb & Horst, 2004).

כתגובה למגבלות אלה התפתחה שיטה חלופית המתבססת על שיפוט מורים. בשיטה זו מורים מנוסים מדרגים את רמת התחכום של תבניות מילים (צורות יסוד) המופיעות בקורפוס כלשהו. במחקר אחד בשיטה זו (Daller et al., 2003), שבעה מורים לטורקית כשפה זרה דירגו 693 מילים לשלוש רמות תחכום לקסיקלי: בסיסיות, בינוניות ומתקדמות, בהתבסס על הניסיון שלהם בהוראה. הממצאים הראו כי שיטה זו יעילה להערכת התקדמות באוצר מילים. באופן דומה במחקר על צרפתית כ-L2 (Tidball & Treffers-Daller, 2008), שלושה מורים דירגו 932 מילים לפי שבע דרגות תחכום לקסיקלי. נמצא כי לבדיקת תחכום לקסיקלי, שיפוט מורים מדויק יותר משימוש ברשימות שכיחות. ממצאים אלה תומכים ביתרונותיה של הגישה האיכותנית להערכת תחכום לקסיקלי.

מגבלות השיטות הקודמות להערכת תחכום לקסיקלי

לשיטה המבוססת על רשימות שכיחות מילים מספר מגבלות מהותיות. ראשית, מידת השכיחות של מילים עשויה להשתנות באופן משמעותי בין קורפוסים שונים (Abu-Rabiah, 2024). מילה הנחשבת שכיחה בקורפוס אקדמי עלולה להיות נדירה בקורפוס עיתונאי ולהפך. שנית, רשימות שכיחות אינן מבחינות בין רכישת שפה כ-L1 לבין רכישתה כ-L2, והרי קיימים הבדלים מהותיים בין שני התהליכים האלה (Abu-Rabiah, 2024). מילה הנרכשת בשלב בסיסי ב-L1 עשויה להירכש בשלב מתקדם יותר ב-L2 וההפך, וזאת בשל צורכי הלומדים השונים וכן הבדלי הגילים בין אוכלוסיות הלומדים. לדוגמה, מילים מתחום הגינון כגון 'חרמש', 'מכוש' ו'מריצה' נרכשות בקרב דוברי L1 באופן טבעי בילדות, אך אצל דוברי L2 הן עשויות להירכש רק בשלבים מתקדמים או לא להירכש כלל, אלא אם כן הלומד נחשף להקשר מקצועי ספציפי.

גם לשיטה המבוססת על שיפוט מורים לרמת התחכום הלקסיקלי של תבניות מילים המופיעות בקורפוס מסוים ישנן מגבלות בולטות. ראשית, נטיות שונות של אותה תבנית עשויות להיות בעלות רמות תחכום לקסיקלי שונות. למשל, צורות הנטייה של תבנית מסוימת

אבורביעה, א' (2023). קוגנטים וקוגנטים כוזבים ככלי להרחבת אוצר המילים ברכישת עברית כשפה שנייה לדוברי ערבית. **לקסיקון**, 19, 32-28. <https://doi.org/10.54301/APQU6073>

אתר אצטבא: **קהילה לומדת עברית**. <https://sites.google.com/etz.tzafonet.org.il/estaba>

Abu-Rabiah, E. (2020). Lexical measures for testing progress in Hebrew as Arab students' L2. *Journal of Language and Linguistic Studies*, 16(3), 1096-1114. <https://doi.org/10.17263/jlls.803551>

Abu-Rabiah, E. (2023). Evaluating L2 vocabulary development features using lexical density and lexical diversity measures. *LLT Journal: A Journal on Language and Language Learning*, 26(1), 168-182. <https://doi.org/10.24071/llt.v26i1.5841>

Abu-Rabiah, E. (2024). Exploring lexical sophistication in second language: An analysis of vocabulary using a word-rating method. *Theory and Practice of Second Language Acquisition*, 10(2), 1-24. <https://doi.org/10.31261/TAPSLA.15623>

Abu-Rabiah, E. (2025). The productive vocabulary size of second language learners upon entry into higher education. *Theory and Practice of Second Language Acquisition*, 11(1), 1-19. <https://doi.org/10.31261/TAPSLA.16594>

Cobb, T., & Horst, M. (2004). Is there room for an academic word list in French? In P. Bogaards & B. Laufer (eds.), *Vocabulary in a second language: Selection, acquisition, and testing* (pp. 15-38). John Benjamins. <https://doi.org/10.1075/llt.10.04cob>

Daller, H., Van Hout, R., & Treffers-Daller, J. (2003). Lexical richness in the spontaneous speech of bilinguals. *Applied Linguistics*, 24(2), 197-222. <https://doi.org/10.1093/applin/24.2.197>

Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24(2), 143-188. <https://doi.org/10.1017/S0272263102002024>

Ellis, R. (2015). *Understanding Second language Acquisition*. Oxford University Press.

Horst, M., & Collins, L. (2006). From faible to strong: How does their vocabulary grow?. *Canadian Modern Language Review*, 63(1), 83-106. <https://doi.org/10.3138/cmlr.63.1.83>

Jarvis, S. (2017). Grounding lexical diversity in human judgments. *Language Testing*, 34(4), 537-553. <https://doi.org/10.1177/0265532217710632>

חייבת להתרחש בהקשרים טקסטואליים אותנטיים ומשמעותיים, ולא רק כרשימות נפרדות.

במישור הדידקטי בחירת הטקסטים הלימודיים צריכה לשקף התאמה דינמית בין רמת התחכום הלקסיקלי למיומנות הלומדים. בכיתות המתחילות יש להציג טקסטים בעלי אחוז גבוה של מילים בסיסיות עם שילוב מבוקר של מילים מתקדמות. ככל שהלומדים מתקדמים, יש להגדיל בהדרגה את שיעור המילים המתקדמות תוך שמירה על איזון המאפשר הבנה והטמעה. גישה זו תומכת בעיקרון הפדגוגי של מעבר מהפשוט אל המורכב, ומאפשרת בניית אוצר מילים באופן הדרגתי ומבוקר.

המגמה האחרונה במשרד החינוך, כפי שהיא באה לידי ביטוי בפיתוח רשימות מילים ייעודיות למורי עברית בבתי הספר הערביים (אתר אצטבא), אכן מכירה בחשיבות ההוראה המפורשת של אוצר מילים לפי רשימות מילים. עם זאת, נדרשת התאמה נוספת של חומרים אלה כך שישקפו לא רק שיקולי שכיחות, אלא גם את עקרונות התחכום הלקסיקלי והמאפיינים הייחודיים של רכישת עברית בקרב דוברי ערבית. התאמה כזו תאפשר מיצוי מלא של הפוטנציאל הלימודי הגלום בגישה המדורגת להוראת אוצר מילים.

בהערכת שפה, הן בכתב הן בעל פה, יש לתת משקל מיוחד לשימוש במילים מתקדמות. מעריכים צריכים להבחין בין נבחנים המשתמשים באוצר מילים בסיסי בלבד, לבין אלה המשלבים מילים מתקדמות. שימוש מוגבל במילים מתחכמות עשוי להעיד על אוצר מילים מצומצם ורמת מיומנות לקסיקלית נמוכה, ועל ההערכה לשקף הבדלים אלה.

סיכום

מדד התחכום הלקסיקלי מספק הערכה מעמיקה יותר של מיומנות לקסיקלית בהשוואה למדד הגיוון הלקסיקלי לבד, שכן הוא מתחשב ברמת תחכומן של המילים השונות. הכרה בכך שברכישת L2 קיימות מילים מתקדמות יותר מאחרות וכי תדירות השימוש בהן ורמת הדמיון שלהן ל-L1 משפיעות על רכישתן, מאפשרת למורים לתכנן את הוראת אוצר המילים ולהעריכו באופן יעיל ומותאם יותר ללומדים.

המקורות

אבורביעה, א' (2020). גיוון לקסיקלי: תיאוריה, פרקטיקה ומה שביניהן. **לקסיקון**, 13, 3-5. <https://doi.org/10.54301/xevu1230>

אבורביעה, א' (2022). **רכישת הלקסיקון העברי בקרב תלמידי תיכון דוברי ערבית בנגב: בחינה אמפירית**. [דיסרטציה]. אוניברסיטת בן גוריון בנגב.

Kyle, K., & Crossley, S. (2016). The relationship between lexical sophistication and independent and source-based writing. *Journal of Second Language Writing, 34*, 12–24. <https://doi.org/10.1016/j.jslw.2016.10.003>

Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics, 16*(3), 307–322. <https://doi.org/10.1093/applin/16.3.307>

Lindqvist, C., Bardel, C., & Gudmundson, A. (2011). Lexical richness in the advanced learner's oral production of French and Italian L2. *International Review of Applied Linguistics in Teaching, 49*, 221–240. <https://doi.org/10.1515/iral.2011.013>

Linnarud, M. (1986). *Lexis in Composition: A Performance analysis of Swedish learners' written English*. Liber Forlag. [doi:10.1017/S0272263100000516](https://doi.org/10.1017/S0272263100000516)

Malvern, D., Richards, B., Chipere, N., & Durán, P. (2004). *Lexical diversity and language development: Quantification and assessment*. Palgrave Macmillan. <https://doi.org/10.1057/9780230511804>

Milton, J. (2009). *Measuring Second language Vocabulary Acquisition*. Multilingual Matters. <https://doi.org/10.21832/9781847692092>

Nation, I. S. P. (1990). *Learning and teaching vocabulary*. Newbury House.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.

Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511732942>

Tekmen, E. A. F., & Daloğlu, A. (2006). An investigation of incidental vocabulary acquisition in relation to learner proficiency level and word frequency. *Foreign Language Annals, 39*(2), 220–243. <https://doi.org/10.1111/j.1944-9720.2006.tb02263.x>

Tidball, F. and Treffers-Daller, J. (2008). Analysing lexical richness in French learner language: What frequency lists and teacher judgement can tell us about basic and advanced words. *Journal of French Language Studies, 18*(3), 299–313. <https://doi.org/10.1017/S0959269508003463>

Vermeer, A. (2000). Coming to grips with lexical richness in spontaneous speech data. *Language Testing, 17*(1), 65–83. <https://doi.org/10.1177/026553220001700103>