

רוחב אוצר המילים הפרודוקטיבי בעברית כשפה שנייה בקרב דוברי ערבית

| איהאב ח'ליל אבורביעה*

רוחב לעומת עומק: שני ממדים של ידע לקסיקלי

ביחס להיבטי הידע על אודות המילים, אוצר המילים של הלומד מתחלק לשני סוגים עיקריים: רוחב ועומק. רוחב או גודל של אוצר מילים (vocabulary size or breadth) מתאר את מספר המילים השונות שהלומד יודע – פסיבית או אקטיבית (Schmitt, 2014). זהו מדד כמותי המכמת את ההיקף של הידע הלקסיקלי. רוחב אוצר מילים נמצא במתאם גבוה עם הבנת הנקרא (Li & Kirby, 2015; Qin, 2015), קריאה אקדמית (Qian, 2002), והוא אף מנבא את רמת המיומנות בארבע אופנויות השפה (Milton et al., 2010; Mairano & Santiago, 2020). רוחב אוצר המילים הפסיבי נבדק באמצעות מבחנים שבהם הלומדים מדרגים את רמת ההיכרות שלהם על אודות מספר מילים כלשהו (Nation & Coxhead, 2021). למשל, במבחן סולם ידע אוצר מילים (Vocabulary Knowledge Scale) מדרגים את רמת הידע על סולם בן חמש דרגות. הדרגה הראשונה כוללת ידע ברמת החשיפה למילה בהקשר כלשהו בלי ידיעת משמעותה ואילו הדרגה האחרונה כוללת שליטה מלאה במילה כולל יכולת שיבוצה במשפט (Paribakht & Wesche, 1993). לעומתו, רוחב אוצר המילים הפרודוקטיבי נבדק באמצעות טקסטים כתובים או דבורים שאדם כלשהו מפיק (Nation & Coxhead, 2021). לדוגמה, בהתבסס על כתביו של שייקספיר נמצא כי רוחב אוצר המילים הפרודוקטיבי שלו עומד על כ־20,000 מילים (Craig, 2011).

הסוג השני הוא עומק של אוצר מילים (vocabulary depth) המתאר את היבטי הידע השונים שיש ללומד על כל מילה (Li & Kirby, 2015). המסגרת התיאורטית המקובלת להתייחסות להיבטי ידע אלו היא של ניישן (Nation, 2013), ולפיה לידיעת מילה יש שלושה מרכיבים מרכזיים: צורה, משמעות ושימוש. היבט הצורה כולל את אופן ההגייה והכתיב של המילה וכן את המבנה המורפולוגי שלה – שורש, מוספיות, נטיות וגזירות. היבט המשמעות מתייחס לקשר בין צורת המילה לבין המושג שהיא

מילות מפתח: אוצר מילים, אוצר מילים פסיבי ופרודוקטיבי, עומק ורוחב אוצר מילים, היבטי ידע של מילה, רשימת מילים שכיחות

מבוא: אוצר מילים - מרכיב מפתח ברכישת שפה שנייה

אוצר מילים (vocabulary) או לקסיקון (lexicon) הוא אחת מאבני היסוד המרכזיות ברכישת שפה שנייה (L2), שכן הוא תנאי הכרחי להבנת שפה ולהפקתה (Wilkins, 1972; Milton & Treffers, 2013; Webb & Nation, 2017; Daller, 2013). מחקרים רבים מצביעים על קשר הדוק בין אוצר מילים לבין שליטה בארבע אופנויות השפה – האזנה, דיבור, קריאה וכתיבה – וכן על תרומתו הרבה להצלחה אקדמית (Pikulski & Templeton, 2004; Sedita, 2005; Webb & Nation, 2017).

סוגי אוצר מילים: פסיבי ופרודוקטיבי

ביחס לארבע אופנויות השפה, אוצר המילים של הלומד מתחלק לשני סוגים עיקריים: פסיבי ופרודוקטיבי. אוצר מילים פסיבי (receptive vocabulary) מתייחס למילים שהלומד מזהה בעת קריאה או האזנה. הכרת מילה זו עשויה להתבסס על הכרת משמעותה הכללית וצורתה הדבורה או הכתובה. אוצר מילים פרודוקטיבי (productive vocabulary) כולל את המילים שהלומד מסוגל לשלוף בזמן אמת ולהשתמש בהן בפועל – בכתיבה או בדיבור (Milton, 2009). מחקרים מראים כי אוצר המילים הפרודוקטיבי מהווה בין 40% עד 80% מאוצר המילים הפסיבי (Milton, 2009; Nizonkiza, 2016; Miao, 2019). הפער הזה נובע מכך שזיהוי מילה בהקשר הוא תהליך הדורש ידע מצומצם יחסית על אודות מילה כלשהי בהשוואה להפקה מדויקת שלה, הדורשת ידע תחבירי, פרגמטי, מורפולוגי ועוד.

* 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, ובמיוחד 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>

שהחלו את לימודיהם האקדמיים במוסד להשכלה גבוהה בדרום הארץ. מטרת המחקר היו לא רק להעריך את היקף אוצר המילים הפרודוקטיבי של משתתפי המחקר, אלא גם לבחון את מידת החפיפה בין המילים ששימשו אותם בכתיבה לבין רשימות תדירות של העברית הכתובה והדבורה, וכן לנתח את סוגי המילים שנעשה בהן שימוש.

לצורך כך נאספו 156 חיבורים טיעוניים שכתבו הסטודנטים, כולם בנושא 'האם על הורים לשלם לילדיהם עבור סיוע במטלות הבית'. החיבורים כללו יחד כ-18,000 תיבות אורתוגרפיות. במסגרת תהליך הניתוח בוצעה למטיציה – החזרת כל מילה לצורת הלמה שלה, כלומר, למילת היסוד שלה מאותו חלק דיבר (להרחבה ראו אבורביעה, 2024).

הממצאים הצביעו על כך שרוחב אוצר המילים הפרודוקטיבי של כלל המשתתפים עמד על 1,023 למות (מילות יסוד) בלבד. ממצא זה דומה לממצאי מחקרים על שפתם של לומדי אנגלית כשפה נוספת: למשל, בקרב תלמידי תיכון דוברי ספרדית נמצא כי אוצר המילים הפרודוקטיבי שלהם כולל 948–1,186 מילים (Montero-SaizAja, 2022), ובקרב סטודנטים סינים, המתמחים באנגלית בשנתם השנייה בלימודים – 1,654 מילים, שמהוות כ-44% מאוצר המילים הפסיבי שלהם (Miao, 2019). למרות זאת, ממצא זה של דוברי ערבית נמוך מהצפוי בהתחשב ברקע הלימודי של הסטודנטים: הם למדו עברית במשך 11 שנות לימוד (מכיתה ב עד יב), בהיקף של בין שעתיים לחמש שעות שבועיות – כלומר יותר מ-1,000 שעות הוראה פורמלית. בהקשר זה אוצר מילים פרודוקטיבי של מעט יותר מאלף מילות יסוד מרמז על רכישת פחות ממילת יסוד אחת לכל שעת לימוד, ומעורר שאלות באשר לאפקטיביות ההוראה לשימוש פעיל באוצר מילים.

בהשוואה לרשימות תדירות של העברית הכתובה (heTenTen21 Pinto, 2018; Hebrew Web Corpus 2021): והמדוברת (FDOSH) נמצא ש-50% ממילות היסוד שבקורפוס חפפו את 1,000 מילות היסוד השכיחות ביותר בשפה. רוב המילים שנעשה בהן שימוש היו מילים שכיחות: שמות עצם כלליים (כגון 'דבר', 'בית', 'ילד'), פעלים שכיחים (דוגמת 'עשה', 'רצה', 'יכול'), ותארים פשוטים (כגון 'טוב', 'קשה', 'גדול'). נוכחותן של מילים אקדמיות הייתה מועטה ('השפעה', 'תחום', 'התנגדות', 'לשנות', 'מסוגל'), ומילים רבות שמצופה לראותן בכתיבה טיעונית לא הופיעו כלל (למשל 'גישה', 'היבט', 'אתגר', 'תועלת', 'להסתגל', 'להודות', 'לצפות', 'להקצות').

ממצאים אלו מצביעים על כך שאוצר המילים הפרודוקטיבי העברי של דוברי ערבית בעת כניסתם ללימודים במוסדות להשכלה הגבוהה נשען בעיקר על מילים שכיחות מאוד. היעדר מילים אקדמיות מעיד על פער בין הדרישות הלשוניות של ההקשר האקדמי לבין היכולות הלשוניות בפועל של הלומדים. פער זה עשוי

מייצגת, למשמעויותיה המרובות, לאסוציאציות שהיא מעוררת ולקשרים הסמנטיים שלה עם מילים אחרות. היבט השימוש עוסק במאפייני הדקדוקיים של המילה (כגון חלק הדיבר), בקולוקציות שהיא משתלבת בהן ובתנאים הפרגמטיים השונים לשימוש בה – דוגמת המשלב הלשוני, רמת הרשמיות ותדירותה בטקסטים שונים. כל אחד מן המרכיבים האלה כולל ידע סביל (הבנה) וידע פעיל (הפקה). עומק אוצר המילים חיוני לתקשורת אפקטיבית ב-L2 (Yanagisawa & Webb, 2019). הוא אף מנבא ביצועים בקריאה אקדמית (Qian, 2002), ומוביל לשיפור בכתיבת סיכום או תמצית ובמשימות הסקה (Elmasry, 2012; Zhang & Yang, 2016).

רוחב אוצר מילים בשפת אם ובשפה שנייה היעדים הנדרשים לתפקוד לשוני

מחקרים מראים כי דוברי אנגלית כשפת אם (L1) מגיעים לבגרות עם אוצר מילים של 10,000–11,000 משפחות מילים (Milton & Treffers-Daller, 2013; Brysbaert et al., 2016). משפחת מילים מוגדרת כאוסף מילים השייכות לאותו שורש ומבוססת על אותה מילת יסוד. לעומת זאת, לומדי L2 מגיעים לרוב, לאחר כ-1,000 שעות הוראה, לרוחב מצומצם של 2,000–4,000 משפחות מילים בלבד (Laufer, 2010). רוחב אוצר מילים כזה מאפשר תפקוד בסיסי בשפה – בעיקר בדיבור יום-יומי, אך הוא אינו מספיק להצלחה בלימודים אקדמיים.

מחקרים מצביעים על כך שבשפה האנגלית נדרש רוחב שונה של אוצר מילים כדי להתמודד עם סוגי תקשורת שונים. לתקשורת בסיסית בדיבור יש להכיר בין 2,000 ל-3,000 משפחות מילים (Van Zeeland & Schmitt, 2013), כדי להבין סרטים ותוכניות טלוויזיה, נדרשות בין 6,000 ל-7,000 משפחות מילים (Webb & Rodgers, 2009a, 2009b), ועבור קריאה שוטפת של טקסטים מגוונים נדרשת שליטה רחבה אף יותר – בין 8,000 לבין 9,000 משפחות מילים (Nation, 2006).

על אף העושר הרב של מחקרים על רוחב אוצר מילים, יש למחקרים בתחום זה מספר מגבלות. ראשית, מחקרים רבים עסקו באוצר מילים באנגלית (כ-L1 וכשפה נוספת) ומעט בשפות אחרות כולל עברית. שנית, מרבית המחקרים העריכו אוצר מילים פסיבי ורק בודדים התמקדו באוצר מילים פרודוקטיבי. יתרה מזו, מחקרים ספורים בחנו אוצר מילים של דוברי L2 בעת הכניסה למוסדות להשכלה גבוהה.

רוחב אוצר מילים פרודוקטיבי בעברית כשפה שנייה לדוברי ערבית

במחקר עדכני בחן אבורביעה (Abu-Rabiah, 2025) את רוחב אוצר המילים הפרודוקטיבי בעברית של 156 לומדים דוברי ערבית,

Brysbaert, M., Stevens, M., Mander, P., & Keuleers, E. (2016). How many words do we know? Practical estimates of vocabulary size dependent on word definition, the degree of language input and the participant's age. *Frontiers in Psychology*, 7, 1116, 1–11. <https://doi.org/10.3389/fpsyg.2016.01116>

Craig, H. (2011). Shakespeare's vocabulary: Myth and reality. *Shakespeare Quarterly*, 62(1), 53–74. <https://doi.org/10.1353/shq.2011.0002>

Elmasry, H. I. (2012). *Depth and breadth of vocabulary knowledge: Assessing their roles in reading comprehension of high-school EFL learners in the UAE* (Master's thesis). The British University in Dubai.

Hebrew Web Corpus 2021: heTenTen21 corpus: <https://www.sketchengine.eu/hetenten-hebrew-corpus/>

Laufer B. (2010). Form focused instruction in second language vocabulary learning. In Chacón-Beltrán R., Abello-Contesse C., Torreblanca-López M.M., López-Jiménez M.D. (eds.), *Further insights into non-native vocabulary teaching and learning* (pp. 15–27). Bristol: Multilingual Matters. <https://doi.org/10.21832/9781847692900-003>

Li, M., & Kirby, J. R. (2015). The effects of vocabulary breadth and depth on English reading. *Applied Linguistics*, 36(5), 611–634. <https://doi.org/10.1093/applin/amu007>

Mairano, P., & Santiago, F. (2020). What vocabulary size tells us about pronunciation skills: Issues in assessing L2 learners. *Journal of French Language Studies*, 30(2), 141–160. <https://doi.org/10.1017/S0959269520000010>

Miao, J. (2019). *A study on English majors' vocabulary size and its correlation with scores on test for English majors grade-four*. In the 2nd International Conference on Contemporary Education, Social Sciences and Ecological Studies (CESSSES 2019) (pp. 202–208). Atlantis Press. <https://doi.org/10.2991/cessses-19.2019.48>

Milton, J. (2009). *Measuring second language vocabulary acquisition*. Multilingual Matters. <https://doi.org/10.21832/9781847692092>

Milton, J., & Treffers-Daller, J. (2013). Vocabulary size revisited: The link between vocabulary size and academic achievement. *Applied Linguistics Review*, 4, 151–172. <https://doi.org/10.1515/applirev-2013-0007>

Milton, J., Wade, J., & Hopkins, N. (2010). Aural word recognition and oral competence in English as a Foreign Language. In R. Chacón-Beltrán, C. Abello-Contesse & M. Torreblanca-López (Eds.), *Insights into non-native vocabulary teaching and learning* (83–98). Multilingual Matters. <https://doi.org/10.21832/9781847692900-007>

להקשות על השתלבותם וכן על הצלחתם במסגרות אקדמיות, ולפיכך יש להקדיש תשומת לב ייחודית להוראה של אוצר מילים אקדמי פרודוקטיבי באוכלוסייה זו ולפיתוחו.

מסקנות

היעדר שימוש במילים אקדמיות עלול להקשות על הלומדים לא רק בהבנת חומרים כתובים והרצאות באקדמיה, אלא גם בהתבטאות בעל פה ובכתב במהלך לימודיהם. יתרה מזו, אף על פי שאוצר מילים פרודוקטיבי של כ־1,000 מילות יסוד תואם ממצאים בקרב לומדי שפה נוספת במדינות אחרות, מדובר בהישג מצומצם יחסית, בהתחשב במספר שנות הלימוד של אוכלוסיית המחקר. ייתכן שהסיבה לכך היא הוראה המתמקדת בתרגום לצורך הבנה בלי עידוד של שימוש פעיל במילים.

יישומים פדגוגיים

בדומה למרתף גדוש כלי עבודה משובחים, אוצר מילים פסיבי הוא הבסיס הקוגניטיבי (ההכרני) האיתן להבנת השפה. אולם, ללא היכולת המעשית להוציא את הכלים מהמרתף ולרתום אותם למלאכה מגוונת – זוהי היכולת הפרודוקטיבית – האוצר הלשוני העצום נותר כפוטנציאל בלבד, טמון ובלתי מנוצל. פיתוח מיומן של אוצר המילים הפרודוקטיבי הוא המפתח להשתתפות פעילה ומשמעותית בחיים בכלל ובלומדים האקדמיים בפרט, הן בדיבור הן בכתובה.

כדי להרחיב את אוצר המילים הפרודוקטיבי של הלומדים יש לעודד הוראה מבוססת שימוש, הכוללת חשיפה למשימות המחייבות כתיבה ודיבור בהקשרים בעלי משמעות עבור הלומדים. כמו כן נוסף להרחבת אוצר המילים יש לטפח גם את עומקו – באמצעות הכרת קולוקציות והבנה של ניואנסים תחביריים ופרגמטיים. יתר על כן, חשוב לשלב הוראה מפורשת של אוצר מילים אקדמי כבר בשלב התיכון, כהכנה ללימודים אקדמיים. כדי להבטיח יעילות בהוראה מומלץ להסתמך, ככל האפשר, על רשימות שכוחות מעודכנות של מילים, המותאמות לדוברי עברית כ־L2.

המקורות

אבורביעה, א' (2024). הגדרות של מילה בהוראת אוצר מילים ובהערכתו: תיבה, תמנית, תבנית, למה, פלמה ומשפחת מילים. *לקסייקיי*, 21, 1–8.

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>

- Webb, S., & Nation, I. S. P. (2017). *How vocabulary is learned*. Oxford University Press.
- Webb, S., & Rodgers, M. P. (2009). The lexical coverage of movies. *Applied Linguistics*, 30(3), 407–427. <https://doi.org/10.1093/applin/amp010>
- Webb, S., & Rodgers, M. P. (2009) b. Vocabulary demands of television programs. *Language Learning*, 59(2), 335–366. <https://doi.org/10.1111/j.1467-9922.2009.00509.x>
- Wilkins, D. A. (1972). *Linguistics in language teaching*. Edward Arnold.
- Yanagisawa, A., & Webb, S. (2019). Measuring depth of vocabulary knowledge. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 371–386). New York, NY: Routledge. <https://doi.org/10.4324/9780429291586-24>
- Zhang, D., & Yang, X. (2016). Chinese L2 learners' depth of vocabulary knowledge and its role in reading comprehension. *Foreign Language Annals*, 49(4), 699–715. <https://doi.org/10.1111/flan.12225>
- Montero-SaizAja, A. (2022). The productive vocabulary of multimodal and unimodal English as a Foreign Language learners. *Porta Linguarum: Revista Internacional de Didáctica de Las Lenguas Extranjeras. Monográfico 4*, 139–153. <http://dx.doi.org/10.30827/portalin.vi.21389>
- Nation, I. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63(1), 59–82. <https://doi.org/10.3138/cmlr.63.1.59>
- Nation, I. S. (2013). *Teaching & learning vocabulary*. Boston: Heinle Cengage Learning.
- Nation, I. S. P., & Coxhead, A. (2021). *Measuring native-speaker vocabulary size*. John Benjamins Publishing Company.
- Nizonkiza, D. (2016). First-year university students' receptive and productive use of academic vocabulary. *Stellenbosch Papers in Linguistics*, 45, 169–187.
- Paribakht, T. S., & Wesche, M. B. (1993). Reading comprehension and second language development in a comprehension-based ESL program. *TESL Canada Journal*, 11(1), 9–29. <https://doi.org/10.18806/tesl.v11i1.623>
- Pikulski, J., & Templeton, S. (2004). *Teaching and developing vocabulary: Key to long-term reading success*. Houghton Mifflin.
- Pinto, J. D. (2018). *Creating a frequency dictionary of spoken Hebrew: A reproducible use of technology to overcome scarcity of data* (Doctoral dissertation). University of Texas at Austin.
- Qian, D. D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: An assessment perspective. *Language Learning*, 52(3), 513–536. <https://doi.org/10.1111/1467-9922.00193>
- Qin, C. (2015). *Didactic considerations of vocabulary breadth and depth in EFL/ESL contexts- a literature review*. Göteborgs University. https://gupea.ub.gu.se/bitstream/2077/38777/1/gupea_2077_38777_1.pdf
- Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning*, 64(4), 913–951. <https://doi.org/10.1111/lang.12077>
- Sedita, J. (2005). Effective vocabulary instruction. *Insights on Learning Disabilities*, 2(1), 33–45. <https://www.cgcatogo.com/uploads/1/0/6/7/10675379/effective-vocabulary-instruction.pdf>
- Van Zeeland, H., & Schmitt, N. (2013). Lexical coverage in L1 and L2 listening comprehension: The same or different from reading comprehension? *Applied Linguistics*, 34(4), 457–479. <https://doi.org/10.1093/applin/ams074>