

גיוון לקסיקלי: תיאוריה, פרקטיקה ומה שביניהן

איהאב ח'ליל אבו רביעה*

מילות מפתח: רכישת שפה שנייה, אוצר מילים, הערכת כתיבה, מדדים לקסיקליים, עברית-ערבית

כיום, רכישת שפה שנייה מהווה מרכיב עיקרי בתוכניות הלימודים במדינות רבות, ולכן יש חשיבות רבה לחקר התחום. תוכניות הלימודים בבתי הספר בישראל מדגישות את חשיבות העברית לכל אזרח במדינת ישראל – הן כשפה ראשונה הן כשפה שנייה. העברית היא שפת מוסדות המדינה, מוסדות האקדמיה, התקשורת והרחוב, ולכן חשוב שכלל האזרחים, כולל האזרחים ששפת אימם ערבית ישלטו בה. העברית היא מפתח להשתלבותם הכלכלית והחברתית של האזרחים ששפת אימם ערבית, ולכניסתם ללימודים גבוהים (משרד החינוך, 2011).

"גיוון לקסיקלי" הנדון במאמר זה הוא מונח מתחום הלקסיקון (אוצר המילים). חקר המילים הוא אחד מענפיה המגוונים של רכישת שפה, לצד דקדוק, תחביר, מורפולוגיה ועוד. כעת, לאחר תקופה ארוכה של התמקדות בחקר הדקדוק (ותחומי שפה אחרים), עולה ההכרה בחשיבות לימוד אוצר המילים, ונבחנות אפשרויות לקידום הוראתו של אוצר המילים (Read, 2000: 2). לגיוון אוצר המילים ועושרו בתוכניות לימודי שפה, יש חשיבות מכרעת הן בכתיבה הן בדיבור, אך עדיין לא הוצעה דרך מדויקת למדידתו.

המילים הן אבני היסוד של השפה, ומהן נוצרים מבנים גדולים יותר, כמו משפטים, פסקות וטקסטים. לומדים רבים רואים ברכישה של אוצר מילים את עיקרה של רכישת השפה השנייה, ולכן הם מקדישים את מרבית זמנם לשינון רשימות מילים. ניתן אפוא לראות ברכישת הלקסיקון עדיפות עליונה בהוראת שפה בכלל ושפה שנייה בפרט, ולכן היכולת להעריך את מידת השליטה באוצר המילים היא חיונית. כדי להעריך עד כמה ההתקדמות ברכישה לקסיקלית משפיעה על כישורי התקשורת של הלומדים הן בכתיבה הן בדיבור, יש להקפיד על הערכת הרכישה בצורה מהימנה (Lu, 2012:190; Read, 2000: 3).

שליטה בהיבטיו השונים של הלקסיקון משלבת שני סוגים של כשירות: כשירות בין-לקסיקלית, כלומר היכולת לבחור מילה נכונה מבין אחרות הדומות לה; וכשירות תוך-לקסיקלית, כלומר היכולת להשתמש במילה אחת בכל ההקשרים הנכונים (Tanaka,

* איהאב ח'ליל אבו רביעה הוא דוקטורנט במחלקה ללשון העברית באוניברסיטת בן-גוריון בנגב בהנחייתם של פרופ' רוני הנקין-רויטפרב וד"ר רועי גפטר. הוא מלמד במרכז ללימודים קדם-אקדמיים באוניברסיטת בן-גוריון בנגב ובהתמחות בלשון העברית במכללה האקדמית לחינוך ע"ש קיי. ערך זה נכתב בתמיכת מלגת נגב – רוח דרומית.

2: 2012). שליטה כפולה זו זכתה לכינויים שונים בספרות: כשירות לקסיקלית, מיומנות לקסיקלית, עומק לקסיקלי ובעיקר עושר לקסיקלי (lexical richness) (Koizumi, 2012: 555).

כדי להעריך באופן כמותי את השליטה הזו פותחו מדדים אחדים. המפורסם מביניהם הוא הגיוון הלקסיקלי (lexical diversity/ lexical variation/ lexical range). מדד זה כשמו כן הוא, מדד המתייחס לטווח אוצר המילים של הלומד כפי שהוא מוצג בשימושו בשפתו (Malvern et al., 2004).

הגיוון הלקסיקלי משפיע על מהימנות הדובר בעיני זולתו, על האופן שבו הנמענים יקבלו אותו ועל היעילות התקשורתית בינו לבין נמעניו (Dillard & Pfau, 2002). לקסיקון מורכב, בהנחה שהוא מוסר מסר מורכב, נתפס כמעניין ומושך, בעיקר משום שהוא מייצג בחירה לקסיקלית מורכבת, המיוחסת לידע ולאמינות. מכאן החשיבות של מדידת יכולת זו וטיפול כמדד ליכולת לשונית ותקשורתית.

הדרך הנפוצה ביותר לחישוב מדד זה היא TTR type-token ratio. נוסחה זו, שהציג לראשונה נדל ג'ונסון (Johnson, 1939), בודקת את הגיוון הלקסיקלי בטקסט באמצעות ספירת התבניות (types) – המילים השונות, חלקי כלל התבניות (tokens) של המילים בטקסט. תבנית היא ערך מילוני על נטיותיו השונות. למשל, המילים: "ילדים" ו"ילדות" נספרות כשתי תבניות של תבנית אחת – התבנית "ילד"; במשפט: "המורה שיבח את התלמיד ואת הוריו" – יש 6 תבניות ו-5 תבניות ("את" מופיעה פעמיים). ככל שמספר החזרות על תבנית קטן, הגיוון הלקסיקלי עולה (Lieven, 1978; Bates, Bretherton & Snyder, 1991). גיוון לקסיקלי גבוה מצביע על אוצר מילים עשיר של הדובר או של הכותב, שכן בשל גיוון זה אין הוא צריך להשתמש בתבניות רבות של כל תבנית (Johansson, 2009: 62).

מחקרים שונים הצביעו על חיסרון מתודי במדד TTR : הוא מושפע מאורך הטקסט (Hess, Sefton & Landry, 1986; Jarvis, 2002, 2013; Malvern et al., 2004; McCarthy & Jarvis, 2007). ההסבר לירידה זו בערכי TTR בגין אורך טקסט הוא שהתבניות של האדם מוגבלות ואילו תבניותיו אינן מוגבלות. כלומר, לאדם יש מספר מסוים של מילים שונות (תבניות) זמינות על נושא כלשהו, אבל הוא יכול לכתוב או להגיד אין-סוף מילים (תבניות) על אותו נושא. יתר על כן, לעיתים קרובות הכותב או הדובר חוזר על תבניות כלשהן לשם חיזוק הטענה או לצרכים פרגמטיים אחרים, כמו הדגשת הרעיון. מחקרים הראו שעם העלייה בגיל הנבדקים (עד לרמת כשירות מסוימת), TTR דווקא יורד (Johansson, 2009, p.63). בשל העובדה שאורך הטקסטים עולים עם הגיל.

בכל מבחני הכתיבה ומטלות הכתיבה בכל רמות ההשכלה הן בשפה ראשונה הן בשפה שנייה.

שילוב בין התיאוריה לפרקטיקה

כאמור, לאוצר מילים מגוון ועשיר חשיבות רבה במבחני השפה השונים, אך אין הגדרה פרקטית לאופן שבו הבדוק יכול להעריך אותו. המורים תופסים באופן שונה גיוון לקסיקלי ועושר לקסיקלי, מפני שהם מלמדים אוכלוסיות שונות בעלות יכולות לשוניות שונות. נוסף על כך, מורים הנבדלים זה מזה בכשירות הלשונית שלהם מעריכים באופן שונה את אותם הטקסטים. חשיפת המורים למחקרים הלשוניים על הערכת אוצר המילים, ובעיקר למחקרים העוסקים בגיוון לקסיקלי, עשויה להקל עליהם משמעותית את המלאכה. כדי להתגבר על כל המשתנים האלה ניתן להשתמש במחקרים הלשוניים לתיעול ולשיקום פרקטיקת הערכת אוצר המילים בכתיבה. למשל, המורים יכולים לבדוק את גיוון אוצר המילים בכתיבה של התלמידים או של הסטודנטים באמצעות מדד ה-TTR רק בשלוש השורות הראשונות. כך יהיו להם נתונים כמותיים מדויקים ובאמצעותם הם יוכלו לתת ציון מדויק יותר לגיוון אוצר המילים של הכותב.

חשיפת המורים למחקרים לשוניים על הערכת אוצר המילים יכולה גם לסייע לתהליכי המחשת מושגים מופשטים. בעקבות כך בדיקת מושגים אלה, שעד כה הייתה שונה מבדוק לבדוק, אמורפית ותלויה באינטואיציה אישית, תהפוך לברורה, לממוסדת ולבלתי-תלויה בגחמות אישיות. יתרה מזו, הבנת ההיגיון העומד מאחורי בדיקת גיוון לקסיקלי יכולה להיטיב את ההערכה של הכתיבה, שכן הבדוק נהיה מודע לגורמים המשפיעים על הבדיקה, כגון אורך הכתיבה, ומתחשב בהם.

תכנון הלימודים בשפה ראשונה וגם בשפה שנייה עובר שינויים בעקבות התפתחויות בתחומים משיקים: בלשנות שימושית, חקר דרכי רכישת שפה, מתודיקה של הוראת לשון ועוד (אולשטיין, 1992). על כן, הנגשת המושג הבלשני "גיוון לקסיקלי" לכל העוסקים בהוראת והערכת שפה ראשונה או שנייה בכל תקופות הגיל ורמות ההשכלה והבהרת דרכי השימוש בו ככלי יעיל להערכת ההתקדמות ברכישת הלקסיקון – המהווה בחלק הארי מן המקרים סממן להתקדמות בכשירות הלשונית הכללית – הן צעד חיוני לקידום החינוך הלשוני.

המקורות

אבו רביעה, א' (2017). **הגיוון לקסיקלי, הדחיסות הלקסיקלית ורמת המופשטות כמדדי התקדמות ברכישת העברית כ-L2 לערבית של צעירים בדואים בנגב**. חיבור לשם קבלת תואר מוסמך בלשון עברית. באר שבע: אוניברסיטת בן גוריון בנגב.

חיסרון זה העסיק חוקרים רבים, והם ניסו להציע פתרונות כדי להתגבר עליו. למשל, בגישת "אוצר המילים התיאורטי" (theoretical vocabulary) מוצעת בחירה רנדומאלית של מספר מילים מהטקסט ברצף או לא ברצף, מיון לתבניות השונות וחישוב הגיוון הלקסיקלי. כך ניתן למדוד טקסטים באורכים שונים מבלי שאורכם ישפיע על החישוב (Johansson, 2009: 65).

מדד הגיוון הלקסיקלי שימש במאות מחקרים על הערכת רכישת הלקסיקון של לומדי שפות שונות בעולם (כמו Kalantari & Gholami, 2017; Zareva, 2019), אך למיטב ידיעתי אף לא במחקר אחד על רכישת העברית כשפה ראשונה או שנייה. בעברית מדד זה נבדק רק כסממן של ז'אנרים שונים בלשון העיתונות אצל כותבים ילידיים של השפה העברית (שלזינגר, 2000). בשל מחסור זה ביצעתי מחקר ראשוני בתחום הרכישה הלקסיקלית. בדקתי שלושה מדדי עושר לקסיקלי, כולל מדד הגיוון הלקסיקלי בעברית כשפה שנייה אצל תלמידים דוברי ערבית כשפת אם, והגעתי למספר תובנות על אודות הוראת העברית לדוברי הערבית בישראל (אבו רביעה, 2017).

הפרקטיקה

המורים המלמדים בבתי ספר מעריכים לעיתים קרובות את כתיבתם של התלמידים (כולל הלקסיקון) הן במבחנים הן בעבודות כתובות. כך גם מרצים המלמדים שפות במכללות ובאוניברסיטאות. בכל מבחני הבגרות בעברית של דוברים ילידים ושל דוברי שפה שנייה יש שאלות על אוצר המילים ושאלות כתיבה שבהן נבדק הלקסיקון. רבע ממבחן הבגרות (החובה) בעברית לדוברי ערבית עוסק בכתיבה. משרד החינוך מספק מחוונים כלליים לבדיקת סוגים שונים של כתיבה ובהם ניתן מקום מרכזי להערכת המבע, הכולל אוצר מילים וביטויים עשיר ומדויק (הערכת כתיבה, 2009).

אוצר המילים אינו חשוב רק אצל תלמידים, אלא גם אצל סטודנטים. מרבית המוסדות להשכלה גבוהה בארץ דורשים ציון בבחינת יע"ל, בחינת ידע בעברית שנועדה לבחון את רמת השליטה של הנבחן בשפה העברית כתנאי קבלה למוסד. למבחן זה ניגשים אלו שנבחנו בבחינה הפסיכומטרית בשפה שאינה עברית ואלו שלא נדרשו להיבחן בבחינה הפסיכומטרית, אך נבחנו בבחינות הבגרות בשפה שאינה עברית. בבחינה זו הכתיבה מהווה שליש מציון המבחן והיא מוערכת לפי ארבעה קריטריונים: תוכן, ארגון, עושר הלשון ודיוק הלשון (מאל"ו 1). כלומר, רבע מציון הכתיבה ניתן על לשון עשירה. מטלת הכתיבה בפסיכומטרי של דוברי העברית מהווה גם היא רבע מהציון הכללי, והיא מוערכת על פי שני ממדים: תוכן ולשון. ציון טוב בלשון כולל בין השאר לשון עשירה, מדויקת וקולחת ושימוש סמנטי מדויק (מאל"ו 2). הווי אומר, גיוון לקסיקלי חשוב מאוד

Johansson, V. (2009). Lexical diversity and lexical density in speech and writing: A developmental perspective. *Lund Working Papers in Linguistics*, 53, 61-79.

Johnson, W. (1939). *Language and Speech Hygiene: An Application of General Semantics*. Institute of General Semantics, Chicago.

Kalantari, R., & Gholami, J. (2017). Lexical complexity development from Dynamic Systems Theory perspective: Lexical Density, diversity, and sophistication. *International Journal of Instruction*, 10(4), 1-18.

Koizumi, R. (2012). Relationships between text length and lexical diversity measures: Can we use short texts of less than 100 tokens?. *Vocabulary Learning and Instruction*, 1(1), 60-69.

Lieven, E. V. M. (1978). Conversations between mothers and Young children: Individual differences and their implications for the study of language learning. In N. Waterson & C. Snow (Eds.), *The Development of Communication* (pp.173-187). London: John Wiley.

Lu, X. (2012). The relationship of lexical richness to the quality of ESL learners' oral narratives. *The Modern Language Journal*, 96(2), 190-208.

Malvern, D., Richards, B., Chipere, N., & Durán, P. (2004). *Lexical diversity and language development*. New York: Palgrave Macmillan.

McCarthy, P. M., & Jarvis, S. (2007). Vocab: A theoretical and empirical evaluation. *Language Testing*, 24(4), 459-488.

Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.

Tanaka, S. (2012). New directions in L2 lexical development. *Vocabulary Learning and Instruction*, 1(1), 1-9.

Zareva, A. (2019). Lexical complexity of academic presentations. *Journal of Second Language Studies*, 2(1), 71-92.

אולשטין, ע' (1992). תכנון לימודים ורכישת שפה שנייה בחברת הגירה. **הלכה למעשה בתכנון לימודים**, 7: 7-15.

הערכת כתיבה - תהליך ותוצר. (2009). נדלה מתוך

http://cms.education.gov.il/EducationCMS/UNITs/Tochniyot_Limudim/Ivrit/Templates/RegularText.aspx?NRMODE=Published&NRNODEGUID=%7b486B622B-A660-4916-BB36-200AD6101D30%7d&NRORIGINALURL=%2fEducationCMS%2fUnits%2fTochniyot_Limudim%2fIvrit%2fDigmey%2fHabaa%2fHaarachatKtiva%2ehtm&NRCACHEHINT=NoModifyGuest#back

מאלי"ו (מרכז ארצי לבחינות ולהערכה) 1. **חוברת הדרכה - בחינת מיון בעברית (יע"ל)**. נדלה מתוך <https://www.nite.org.il/files/yael/examples/yael-hebrew.pdf>

מאלי"ו (מרכז ארצי לבחינות ולהערכה) 2. **חוברת הדרכה - בחינת הכניסה הפסיכומטרית לאוניברסיטאות**. נדלה מתוך https://nite.org.il/files/psych/new_psych/verbal-hebrew.pdf

משרד החינוך, (2011). **עברית כשפה שנייה בבתי הספר הערביים לשלוש חטיבות הגיל: ג'-י"ב**. נדלה מתוך <http://cms.education.gov.il/NR/rdonlyres/202A56F5-EC9E-43CD-9592-0EAE76451DF2/212803/29916.pdf>

שלזינגר, י' (2000). **לשונות העיתון: מאפיינים סגנוניים בלשון העיתונות היומית הכתובה על מדוריה השונים**. באר-שבוע: הוצאת ספרים של אוניברסיטת בן-גוריון בנגב.

Bates, E., Bretherton, I., & Snyder, L. S. (1991). From first words to grammar: *Individual differences and dissociable mechanisms* (Vol. 20). Cambridge University Press.

Dillard, J. P., & Pfau, M. (2002). *The persuasion handbook: Developments in theory and practice*. Sage Publications.

Hess, C. W., Sefton, K. M., & Landry, R. G. (1986). Sample size and type-token ratios for oral language of preschool children. *Journal of Speech, Language, and Hearing Research*, 29(1), 129-134.

Jarvis, S. (2002). Short texts, best-fitting curves and new measures of lexical diversity. *Language Testing*, 19(1), 57-84.

Jarvis, S. (2013). Capturing the diversity in lexical diversity. *Language Learning*, 63, 87-106.