

## עמליה בר-און

### קריאה בצל ההומוגרפיות: הבעיה היא גם הפתרון

כרבע מהמילים בטקסט העברי הלא מנוקד הן מילים הומוגרפיות, כלומר מילים שאפשר לפענח בדרכים שונות. את רצף האותיות **הפרה**, כדוגמה, אפשר לפענח כשם הפעולה **הִפְרָה**, כשם עצם מיודע **הִפְרָה**, כמוספית תחילית – **הִפְרָה** היסטורי, כפעלים בצורת הנקבה – **הִפְרָה** ובצורת הזכר – **הִפְרָה** וכפועל בצורה המודלית – **הִפְרָה**. ריבוי המילים ההומוגרפיות מכתוב קריאה תלוית הקשר. ההקשר הוא חלק בלתי נפרד מזיהוי המילה העברית בטקסט, והודות לו מצליחים הקוראים המיומנים לקרוא באופן שוטף ומדויק במערכת כתב עמומה. עם זאת, ההקשר אינו ערובה מלאה לפתרון ההומוגרפיות; כך למשל ייתכן כי שטף הקריאה יִפְגַע בפענוח חלקו הראשון של המשפט **למרות שהפרה במשפט אינה נמוצה**...: חלק מהקוראים עשויים להתחיל ולקרוא "למרות שֶהִפְרָה", וסביר שהם יתקנו מיידית את טעותם. העובדה כי מקרים אלה אפשריים מן הצד האחד, אך הם חריגים למדי, מן הצד השני, מעלה את השאלה העומדת במרכזו של מאמר זה: **מה הן תכונותיו של ההקשר הגורמות לכך שהוא מספק מענה כמעט מלא לפתרון העמימות בטקסט העברי?** הטענה המוצגת כאן היא כי מרבית המילים ההומוגרפיות מייצגות חלופות בעלות תכונות מורפו-תחביריות שונות, ותכונה זו מכתובה הקשר מורפו-תחבירי המוביל לפענוח של חלופה אחת, ולא מאפשר, או כמעט שאינו מאפשר, את פענוחה של האחרת. לביסוס טענה זו יציג המאמר ניתוח של תכונותיהן הלשוניות של המילים ההומוגרפיות ויגזור מכך ניתוח של סוגי ההקשרים בטקסט העברי.

### פענוח המילה הלא מנוקדת

הכתב העברי העתיק היה במקורו עיצורי, וכבר במאה התשיעית לפני הספירה נוספו לו אותיות אהו"י שייצגו במקצת את התנועות במילה. 1500 שנה לאחר מכן, בתקופה שבה שימשה העברית בעיקר בתפילות ובטקסטים דתיים, עלה הצורך להפיג את העמימות הפונולוגית בקריאה של הטקסטים, ובעקבות זאת נוספו סימני הניקוד. במאה התשע עשרה, עם החייאתה מחדש של העברית כשפה דבורה, חל תהליך טבעי של השמטת הניקוד, והמערכת חזרה להיות עיצורית בעיקרה עם סימון חלקי ועמום של התנועות באמצעות אותיות אהו"י (שמרון, 1997; Ravid, 2005). המערכת הלא מנוקדת היא ברירת המחדל של קוראי העברית המיומנים, ומדהימה עוד יותר העובדה כי כבר בכיתות ג'-ד' זונחים רוכשי הקריאה את הכתב המנוקד השקוף, שדרכו למדו את הקריאה, ומעדיפים על

פניו מערכת עמומה במיוחד (Bar-On & Ravid, 2011). במובן זה מקיימת מערכת הכתב העברית את עקרון 'אופטימיזציית המידע' (Frost, 2012): מערכת המספקת לקורא מידע פונולוגי וסמנטי מרבי תוך שימוש במספר מצומצם של יחידות אורתוגרפיות.

בעברית הלא מנוקדת חסר מידע על העיצורים במילה: האותיות **ב, כ, פ ו-ש** הן הומוגרפיות – את שלוש הראשונות אפשר להגות כפוצצות וכחוככות, ואת האות **ש** כ-*s* וכ-*sh*. ואולם עיקר המידע הפונולוגי החסר קשור לייצוג התנועות במילה; חמש התנועות *a, e, i, u, o* מיוצגות על ידי אותיות אהו"י, שמעצם הגדרתן הן אותיות הומוגרפיות – כל אחת מהן מייצגת עיצור ותנועה. נוסף על כך, כל אחת מהן מייצגת שתי תנועות: ה-*ו* א מייצגת את התנועות *a* ו-*ו* ו את *o* ו-*ו* ו-*י* את התנועה *i* בעיקר, אך גם את *e*. שחזור התנועות במילה הלא מנוקדת מתבסס על הימצאותן או היעדרן של אותיות אהו"י, וכאן יש להפריד בין תנועות הנמצאות בסופה של המילה ובין תנועות פנימיות. ייצוג התנועות בסוף המילה יציב ועקבי ביותר: חמש התנועות תמיד מיוצגות בסוף המילה, למעט שני מקרים חריגים שבהם התנועה *a* לא מיוצגת – צורת הזכר בגוף שני במבנים חבורים (בגללך, אחיד, אהבתך), ובפעלים בזמן עבר בצורת הזכר היחיד בגוף שני (אהבת, דיברת, התקדמת). ייצוג התנועות הפנימיות מתנהג אחרת: התנועות *o* ו-*u* מיוצגות באופן מלא למדי באמצע המילה, ואילו התנועות *a* ו-*e* אינן מיוצגות באמצע המילה, למעט מקרים של **א** שורשית שאינה נהגית כעיצורית (יצאת, יוצאת). ייצוגה של התנועה הפנימית *i* כפוף למגבלות מורפו-פונמיות, המגדירות שלושה מצבים: הקשרים שייצוגה בהם מחויב (מרגיש, ילדים, רעיל), הקשרים שייצוגה בהם הוא "אסור" (התרגש, הרגיש) והקשרים שבהם ייצוגה אופציונאלי ונתון לבחירתו של הכותב (דיבור או דבור) (רביד, 2011; Ravid & Schiff, 2004). המילה מכתב, כדוגמה, נטולת אותיות אהו"י, והיעדרן של אלו מספק מידע רב על התנועות האפשריות בה: המילה מסתיימת בעיצור ולא בתנועה, שלוש האותיות הראשונות יכולות להיקרא בצירוף של התנועות *a* ו-*e* (למשל *maxatav, mexetev, maxetav*), והתנועה *i* יכולה להופיע לאחר האות **מ** בלבד כחלק מהברה סגורה (*mix*). נוסף על כך, אילוצים פונוטקטיים החלים על המילה העברית (שורצולד, 2002) מכתיבים כי מבין שלוש האותיות הראשונות אפשר להגות רק את השנייה ללא תנועה, כלומר כסוגרת הברה, ואת האות **ב** אפשר להגות כחוככת בלבד. הידע על ייצוג התנועות במילה יחד עם המגבלות הפונוטקטיות מצמצם באופן ניכר את אפשרויות פענוחה, אך מותיר עדיין מגוון רחב של צורות פענוח, הכוללות קומבינציות תנועיות שונות, מבני הטעמה שונים, ובדוגמה לעיל גם את פענוחה של האות **כ** כפוצצת או כחוככת.

צמצום נוסף של האפשרויות נעשה באמצעות המגבלות המורפולוגיות והמורפו-פונולוגיות החלות על המילה העברית, ואלה מותירות למשל במילה המודגמת לעיל שלוש צורות אפשריות בלבד: *mixtav, mexatev, maxtev*, המתאימות למשקלים *maCCeC*, *meCaCeC*, *miCCaC*. הצורה הראשונה – מְכַתֵּב – היא מילה קיימת ונפוצה בשפה, השנייה – מְכַתֵּב – היא מילה ששימשה בתנ"ך וחזרה לשימוש בשנים האחרונות בהקשרים של העידן האלקטרוני, והצורה השלישית אינה מילה קיימת בשפה, וייתכן שבבוא היום היא

תיהפך לכוז. הידע על המבנים המורפולוגיים והמורפו-פונולוגיים של המילה העברית הוא כלי פענוח רב עצמה, בעיקר עבור קוראים צעירים, המגיעים אל רכישת הקריאה מצוידים בידע עשיר על מבנה המילה העברית (Berman, 2003), ויכולים לעשות בו שימוש במעבר מן הכתב המנוקד אל הכתב הלא מנוקד. גם קוראים מיומנים נדרשים לא פעם לכלי פענוח אלו בפוגשם מילים חדשות לא מנוקדות, כמו המילים תקליל או ציימנות שהן בעלות מבנה מורפולוגי פנימי (בר-און, 2010; Bar-On & Ravid, 2011). מפגש זה בוודאי אינו נדיר, על רקע הצטרפותן של מילים חדשות רבות לעברית בת זמננו, שהיא לשון עתיקה המתאימה את עצמה למציאות מודרנית (ניר, 1993). ואולם, יש לסייג ולומר כי כלי הפענוח המורפו-פונולוגיים אינם יעילים תמיד, שכן מילים רבות המצטרפות לשפה הן שאלות, כמו המילה פסטורלי, ובמקרים אלה נדרשת היכרות של המילה, או לחילופין ניקודה.

קורא העברית הצעיר לומד אפוא כבר בראשית דרכו לחפש אחר המבנה התנועי המתאים למילה הכתובה. כך הוא מתרגל לדפוס קריאה המאפשר תנועה בין מבנים תנועיים שונים התואמים את אותה היחידה הכתובה, והרגל זה נמשך אל הקריאה הבוגרת. הקורא המיומן אמנם מכיר את מרבית המילים הכתובות, כך שכמעט ואינו נדרש לחיפוש אחר צורת הפענוח הנכונה, ואולם הוא מודע לכך שיחידות כתובות רבות הן הומוגרפיות, כלומר, מילה כתובה אחת מייצגת מבנים פונולוגיים שונים המביעים משמעויות שונות, כמו הדוגמה לעיל – מכתב. המאמר הנוכחי עוסק בקריאה בצל ההומוגרפיות, ומטרתו להעמיק בהבנת האינטראקציה המתקיימת בין המילה ההומוגרפית וההקשר שבו היא נטועה. הטענה המועלית כאן היא כי הבעיה – ההומוגרפיות, היא גם המפתח לפתרון – ההקשר, וכי תכונותיהן הלשוניות של המילים ההומוגרפיות הן אלו האחראיות ליצירתם של הקשרים תחביריים שונים, אשר בזכותם מתאפשרת קריאה רהוטה, יעילה ומוצלחת של מערכת עמומה במיוחד. לביסוס טענה זו אפתח בתיאור הקריאה בצל ההומוגרפיות, אמשיך בתיאור תכונותיהן הלשוניות של המילים ההומוגרפיות בעברית ואסיים בהצגת הדרך שבה תכונות אלו יוצרות סוגים שונים של הקשרים, כאלה המאפשרים את הקריאה השוטפת, אך לפעמים גם מחבלים בה.

### קריאה בצל ההומוגרפיות

ניסוי קטן שנעשה על 24 קוראים בוגרים ומיומנים ממחיש את הקריאה בצל ההומוגרפיות. הנבדקים התפלגו באופן שווה בין בנים ובנות, וטווח גיליהם נע בין 23 ל-55: 15 בגילאי 26-23, שבעה בני 50-55, ושניים בני 30 ו-42. הנבדקים הצעירים היו סטודנטים, הלומדים נושאים שונים כמו משפטים וסוציולוגיה, או כאלה העומדים לקראת לימודים, והבוגרים בעלי השכלה אקדמית – עורכי דין, אנשי כלכלה וכדומה. הנבדקים התבקשו לקרוא בקול רם קטע מתוך העמוד הראשון בספרו של מאיר שלו 'הקבר היה ככה':

... הצחקוקים נעשו צחוקים ואימתו את חדותי. בעודי ישן התגנבו אלי רוני ונעמי, שתי בנותיו הקטנות של אחי, וצבעו את ציפורני רגלי. הצעירה, כך הסבירו השתיים אחר-כך, צבעה ארבע אצבעות, והבכירה שש. הן קיוו שלא אבחין בזה, אצא כך

אל הבריות ואהיה ללעג ולקלס. אבל עכשיו, כשהתגלתה מזימתן, פרצו אל החדר בקריאות: "אל תוריד את הצבע, אל תוריד, זה נורא יפה."

כל הנבדקים קראו את הפסקה באופן רהוט למדי, ואין זה מפתיע היות שמדובר בקוראים בוגרים ומיומנים. ואולם, מעניין היה לגלות כי אצל מרביתם (18 קוראים) הופיעו "חריקות" קלות במהלך הקריאה, שהתבטאו בהתמהמהות קלה לפני קריאת המילים, ובשגיאה אחת או שתיים, שתוקנו בחלק מהמקרים תוך כדי יצירתן. מרבית השגיאות היו קשורות למבנה התנועי של המילה, כאשר התעלמנו מקריאה 'לא תקנית', שהופיעה אצל כל הנבדקים – מבוגרים כצעירים – כמו קריאת המילה נעשו כ- נַעְשׂוּ במקום נַעְשׂוּ, או קריאת המילה בזה כ- בָּזָה במקום בָּזָה. מעט השגיאות שהופיעו היו בהקשר של המילים ההומוגרפיות, אך גם במילים אחרות, והמילים שמשכו את מרב השגיאות היו רגלי, ואימתו (ארבע שגיאות בכל אחת) ו- ולקלס. בשתי הראשונות נדון בהמשך המאמר, ואולם מעניינת במיוחד המילה ולקלס שאֶתגרה בעיקר את הצעירים: שישה מביניהם קראו את המילה בווריאציות תנועיות שונות, כמו *velekalas*, ורק אחת תיקנה לצורה הנכונה. נראה שדוגמה זו ממחישה את התיאור שהוצג לעיל: מפגש של קורא בוגר עם מילה עברית לא מנוקדת שאינה מוכרת לו. האופן שבו קראו הנבדקים את הקטע הקצר מעיד כי הייצוג העמום של התנועות גובה מחיר מסוים משטף הקריאה, גם כאשר מדובר בקוראים בוגרים ומיומנים. ייתכן כי הדרשה לקריאה קולית, המאלצת את הקורא לשחזר את המבנה הפונולוגי המלא של המילה, עודדה וזימנה שגיאות מעין אלו, ונשאלת השאלה עד כמה אפשר להסיק מכך על הקריאה הדמומה, שהיא הקריאה הבוגרת הטבעית (Share, 2008)? בין חוקרי הקריאה אין תמימות דעים לגבי מידת מעורבותה של הפונולוגיה בנגישות למשמעות המילה; פרוסט (Frost, 1998) מציג את הטענה כי נגישות ישירה למשמעות יכולה אולי לחול במילים מוכרות מאוד, ואולם לרוב הפונולוגיה נושאת תפקיד ראשוני בנגישות למשמעות, גם אם לא בלעדי. מידת שחזורה עשויה להיות שונה בין מילים, ולפחות בזהו מילים באנגלית, ככל שהמבנה התנועי של המילה יהיה מורכב יותר, כך אפשר יהיה להסתמך פחות על הזיהוי הראשוני והמהיר של האותיות העיצוריות, ויידרש שחזור ממצה יותר של המבנה התנועי של המילה (Berent & Perfetti, 1995). בהסתמך על זאת אפשר להסיק כי הפונולוגיה מעורבת בקריאה הדמומה אך במידה פחותה ממעורבותה בקריאה הקולית, כך שגם אם אי אפשר לעשות הקבלה מלאה ביניהן, הרי שאפשר ללמוד מהקריאה הקולית על הקריאה הדמומה, ומעניין להמשיך ולבדוק את הקריאה המיומנת והבוגרת של הטקסט הלא מנוקד.

לאחר הקריאה הקולית נשאלו הנבדקים אם הבחינו במהלך קריאתם במילים ההומוגרפיות בטקסט, כלומר מילים שאפשר לקרוא בכמה דרכים, ומחציתם ענו בשלילה. כהמשך לכך הם התבקשו לסמן את כל המילים ההומוגרפיות בקטע, זאת לאחר מתן הסבר קצר והדגמה של מילה ההומוגרפית, והוקצב לכך זמן סביר של דקה. הפסקה מתוך ספרו של מאיר שלו מוצגת להלן שנית, כאשר המילים ההומוגרפיות שבה מסומנות בגופן מודגש, ומילים שמופיעות יותר מפעם אחת מסומנות גם בקו תחתי.

... הצחקוקים נעשו צחוקים ואימתו את חשדותי. בעודי ישן התגנבו אלי רוני ונעמי, שתי בנותיו הקטנות של אחי, וצבעו את ציפורני רגלי. הצעירה, כך הסבירו השתיים אחר-כך, צבעה ארבע אצבעות, והבכירה שש. הן קיוו שלא אבחין בזה, אצא כך אל הבריות ואהיה ללעג ולקלס. אבל עכשיו, כשהתגלתה מזימתן, פרצו אל החדר בקריאות: "אל תוריד את הצבע, אל תוריד, זה נורא יפה."

הטקסט בן 62 המילים (הצירוף אחר-כך נספר כשתי מילים בניסוי זה) כולל 27 מילים הומוגרפיות (44% מהמילים בטקסט) בספירת תמנית, כלומר ספירת כל המופעים של המילים ההומוגרפיות, כולל אלו שחוזרות כמה פעמים, ו-23 מילים הומוגרפיות (37% מהמילים בטקסט) בספירת תבנית. לפחות שלוש מילים מייצגות יותר משתי חלופות (למשל, אָבֵל, אָבֵל, אָבֵל). הספירה לא כוללת את שתי המילים הקטנות ובקריאות המייצגות חלופות נוספות אפשריות אך נדירות למדי (הֶקְטָנוֹת, וּבְקָרְיָאוֹת), וכמו כן, הספירה מתייחסת אך ורק לרצפים הומוגרפיים, ומתעלמת ממילים דו-משמעיות, למשל, שתי המשמעויות של המילה הקטנות. שימרון וסיון (Shimron & Sivan, 1994) בדקו טקסטים שונים (חומר לימודי, מאמרים אקדמיים וסיפורים עיתונאיים) ומצאו שכרבע מהמילים בטקסטים הן הומוגרפיות, ואלון (1995) מציין שכ-40% מהמילים בטקסטים הן הומוגרפיות. תפוצת המילים בטקסט לעיל תואמת את הצעתו של אלון, וגם אם אחוז המילים ההומוגרפיות בטקסטים אחרים נמוך מתפוצתן בקטע זה, אין ספק שהטקסט העברי רווי במילים מעין אלו.

ההשערה שהקדימה את הניסוי הייתה שהנבדקים יזהו ויסמנו את מרב המילים ההומוגרפיות, אך לא את כולן. ואולם, לא כך היה הדבר: מבין 23 המילים סומנו במוצע 4.67 מילים בלבד (סטיית תקן = 3), וטווח הזיהוי נע בין מילה אחת ל-11 מילים. התמונה המפתיעה חזרה על עצמה: הנבדקים ישבו מול הטקסט, קראו אותו בעיון (רובם בקריאה דמומה), והיו "עיוורים" למרבית המילים ההומוגרפיות שבו. דוברי העברית מודעים לתופעת ההומוגרפיות ומכירים אותה היטב. עדות לכך נראתה במחקרה של בר-און (2010), שבדקה העלאה של חלופות שונות לאותה מילה הומוגרפית. במחקר השתתפו 154 תלמידים ממיצב סוציו-אקונומי בינוני-גבוה משש קבוצות גיל: תחילת כיתה ב' (27 תלמידים), סוף כיתה ב' (26 תלמידים), כיתה ג' (26 תלמידים), כיתה ד' (28 תלמידים), כיתה ז' (26 תלמידים) וכיתה י"א (21 תלמידים). כמו כן כלל המחקר 17 מבוגרים אורייניים. 171 הנבדקים התבקשו לפענח מילים הומוגרפיות במרב הדרכים האפשריות, למשל יבשה (יְבֹשָׁה, יְבֹשָׁה, יְבֹשָׁה, יְבֹשָׁה, יְבֹשָׁה) או צבע (צָבַע, צָבַע, צָבַע, צָבַע). החל מכיתה ג' הצליחו מרבית הנבדקים לשלוף לפחות שתי חלופות למילה, ומספר החלופות עלה עם הגיל. הכישלון החרוץ שנחלו הנבדקים בניסוי הנוכחי בזיהוי המילה ההומוגרפית, כלומר, בזיהוי מילים שיש להן חלופה נוספת לזו המופיעה בטקסט, קשור אפוא לעובדה שהם התבקשו לעשות זאת כאשר היא הייתה נתונה בהקשר של הטקסט ולא במבודד.

התמונה שעלתה בניסוי פשוט זה מלמדת על קריאת הטקסט העברי: הטקסט רווי במילים הומוגרפיות, אבל אלה כמעט שאינן פוגעות בתהליך הקריאה. יתרה מכך, לרוב אין הקורא יודע ואינו מבחין כי במהלך קריאתו הוא מתמודד שוב ושוב עם מצבי עמימות ופותר

אותם בהצלחה רבה, וגם כאשר תשומת לבו מופנית למילים אלו, והוא מתבקש לזהותן בטקסט באופן מודע וסדור, הוא מתקשה באיתורן. כל אלה מעידים על עצמתו של ההקשר, ומראים עד כמה הוא בלתי נפרד מזיהוי המילה העברית הכתובה; למעשה, מהרגע שבו המילה מוצבת בהקשר, הקורא כמעט שאינו מצליח לנתח במנותק ממנו – דרישה מעין זו מחייבת הפרה של התהליך הקוגניטיבי האוטומטי והטבעי של קריאת העברית.

תפקידו של ההקשר בזיהוי מילים כתובות נמצא במרכזו של פולמוס מחקרי מזה שנים רבות: מן הצד האחד עומדת תפיסתו של גודמן (Goodman, 1970), המתארת את הקריאה כתהליך אקטיבי, שבו הקורא מתבסס על רמזי ההקשר כדי לזהות את המילה. תפיסה זו מקבלת תמיכה ממחקרים המתארים את תופעת החמצת אותיות במילים שכיחות, כמו מילות פונקציה, והמעידים על תהליכי מעלה-מטה המשתתפים בקריאה, כלומר תהליכי הקשר (Healy, 1994). כנגד אלו נטען כי זיהוי אוטומטי ומהיר של מילים בודדות הוא הבסיס לקריאה יעילה ושוטפת, וכי ההקשר כמעט שאינו נחוץ לקריאה (Breznitz & Share, 1992; Perfetti, 1985, 1992). בעברית, וכך גם בערבית (Saiegh-Abu-Rabia, 2001; Haddad, 2003), אי אפשר לדבר על קריאה נטולת ההקשר, שכן ההומוגרפיה הנרחבת בשתי השפות יוצרת עמימות לקסיקלית, שפתרונה מחייב הישענות על ההקשר. עמימות לקסיקלית נוצרת גם על רקע של מילים הומונימיות – מילים הנהגות ונכתבות בצורה זהה, אך נושאות משמעויות שונות, כמו המילה **bank**, המייצגת שתי משמעויות: בנק ושפת הנחל, או המילה קבלה, המייצגת שלוש משמעויות: שם הפעולה של הפועל קיבל, אישור על קניית דבר-מה ותורת הסוד והנסתר היהודית. במחקר שנעשה על השפה האנגלית נמצא כי 44% מהמילים שנבחרו במדגם מקרי היו בעלות יותר ממשמעות אחת, וכאשר המדגם כלל מילים שכיחות מאוד עלתה תפוצתן ל-85% (Britton, 1978). מדובר אפוא בתפוצה רחבה של מילים עמומות הדורשות אסטרטגיות פענוח תואמות.

שאלת פתרון העמימות הלקסיקלית מעניינת חוקרים רבים, והיא נבדקת במתודולוגיות מחקריות מגוונות, כמו מעקב אחר תנועות עיניים (Kambe, Rayner, & Duffy, 2001), או מחקרי רגישות המיספריאלית (Burgess & Simpson, 1988). הדיון המרכזי במחקרים אלו הוא טיבה של האינטראקציה המתקיימת בין שכיחות המשמעויות שהמילה מייצגת (למשל, המשמעות הדומיננטית במילה קבלה היא אישור על קניית דבר-מה והפחות דומיננטית היא תורת הנסתר), לבין סוג ההקשר שבו המילה מוצבת: ההקשר יכול להיות ניטרלי, כלומר הקשר המאפשר את שתי החלופות ('היא סיפרה לו על הקבלה'), הקשר המכוון לחלופה הדומיננטית ('הלך ביקש מהמוכר קבלה') והקשר המכוון לחלופה הלא דומיננטית ('הסטודנט הדתי לומד קבלה') (Peleg & Eviatar, 2008). המודלים המתארים את פתרון העמימות הלקסיקלית חלוקים ביניהם באשר לדרך שבה מתקיימת האינטראקציה, אבל כולם מייחסים תפקיד מכריע לגורם השכיחות, ומסכימים כי ההקשר נדרש לשם השגת המשמעות המתאימה. תופעת ההומוגרפיות נדירה למדי בשפה האנגלית, ומילים ספורות הן הומוגרפיות, כמו למשל המילה **wind** המייצגת את המילים *wind* (רוח) ו-*waynd* (מתפתל) (Twilley, Dixon, Taylor, & Clark, 1994). המחקר על פתרון העמימות הלקסיקלית בשפה

האנגלית נעשה אפוא ברובו על מילים הומונימיות, והשאלה היא עד כמה ניתן ליישם את המודלים השונים המוצעים לפתרון העמימות הלקסיקלית על פתרון המילים ההומוגרפיות? פלג ואביתר (Peleg & Eviatar, 2008) בדקו רגישות המיספריאלית למילים ההומוגרפיות והומונימיות בעברית, ומצאו שאין הבדל באופן תהלוכה. במחקרן הן השתמשו במילים המייצגות חלופות מאותה הקטגוריה הלקסיקלית, הנבדלות בדרגת שכיחותן, כמו המילה ספר המייצגת את שמות העצם ספר וספר. הטענה שארצה לטעון כאן היא כי ההקבלה בין המחקרים על פתרון העמימות הלקסיקלית לפתרון המילים ההומוגרפיות העבריות אינה מלאה, שכן המאפיין הבולט של המילים ההומוגרפיות הוא שלרוב הן מייצגות חלופות הנבדלות ביניהן בתכונותיהן המורפו-תחביריות, והבדלים אלה מציבים את גורם השכיחות במעמד מרכזי פחות בפענוח המילה. ארצה להמשיך ולטעון כי החלופות של המילה יוצרות סביבן הקשרים הנבדלים בראש ובראשונה בתכונותיהם המורפו-תחביריות, ולהבדלים אלו השפעה מכרעת על פתרון ההומוגרפיות. להלן יובא תיאור התכונות הלשוניות של המילים ההומוגרפיות וכהמשך לכך – תיאור סוגי ההקשרים הנוצרים סביבן.

### התכונות הלשוניות של המילים ההומוגרפיות

סקירה מקיפה אך לא מלאה של המילים ההומוגרפיות העלתה כי אפשר לתאר מקורות שונים להומוגרפיה, ולמפותן לקבוצות על פיהם (בר-און, 2010). התיאור להלן אינו יכול ואינו מתיימר להציג את המגוון הרחב של הצורות ההומוגרפיות, ואולם הוא עשוי לתרום להבנה של האינטראקציה המתקיימת בין המילה ההומוגרפית וההקשר שבו היא נטועה.

#### מקור ראשון להומוגרפיה: אופייה המשורג של המילה העברית

כל הפעלים בעברית ושמות ותארים רבים בנויים מהשתררות של התבנית המורפולוגית, הכוללת תנועות ולעתיים גם עיצורים תחיליים ו/או סופיים, בין עיצורי השורש. כך למשל הצבת התבנית המורפולוגית CaCaCa על עיצורי השורש  $y-v-\dot{s}$  תיצור את המילה יִבְּשָׁה. שורשים רבים מרכזים סביבם משפחות מרובות מילים, הנבדלות ביניהן בתבניתן המורפולוגית (יִבְּשָׁה, מִיִּבְּשָׁה, יִבְּשָׁה, יִבְּשָׁה, מִיִּבְּשָׁה, יִבְּשָׁה, יִבְּשָׁה), ואלהן יכולות להתווסף מוספיות הגזירה הקווית (יִבְּשָׁתִּי, יִבְּשָׁתִּי). התבניות המורפולוגיות כוללות את משקלי השם והתואר ואת שבעת בנייני הפועל, ולכך יש להוסיף את המבנים המסמנים את שינויי הזמן בפועל, שהם בעיקרם משורגים (ראו למשל את חמש הצורות של השורש בר"ח בבניין קל: בָּרַח, בֹּרַח, יִבְּרַח, בָּרַח, לְבָרַח) (רביד, 2010). המידע החשוב לענייננו כאן הוא שתבניות מורפולוגיות רבות נבדלות ביניהן במבני התנועי בלבד. כך למשל משקלי השם והתואר השונים, וכן אחת מצורות הפועל, במבנים התנועיים CaCoC (שְׁעוֹן, חֲלוֹן; גְּדוֹל, מְתוֹק; אֵדָם, עֵמָק), CaCuC (חֲתוּל, פְּדוּר; כְּתוּב, שְׁמוּר, עֲצוּב) ו-CCoC (כְּפוּר, צְרוּר; שְׁמַר, כְּתַב). המפגש של תכונה זו עם הייצוג החסר והעמום של התנועות במילה הכתובה הלא מנוקדת יוצר מבנים מורפו-אורתוגרפיים זהים המייצגים תבניות מורפולוגיות שונות, ומבנים

אלו אחראים ליצירת כמות אדירה של מילים הומוגרפיות. למעשה, אפשר למפות מילים הומוגרפיות רבות על פני המבנים המורפו-אורתוגרפיים שאליהם הן משתייכות. להדגמת המבנים המורפו-אורתוגרפיים נשתמש בשיטת ייצוג שבה אותיות השורש מיוצגות על ידי הספרה 0, ואותיות התבנית – העיצוריות והתנועיות – נכתבות על פי מיקומן ביחס לאותיות השורש. על פי שיטה זו, המבנה המורפו-אורתוגרפי של שלוש התבניות לעיל הוא 0100. מבנה זה ייצור למשל את המילים ההומוגרפיות כתוב (כְּתוּב, כְּתַב), שמור (שְׁמוּר, שְׁמַר), חמור (חְמוּר, חְמוֹר), שעון (שְׁעוֹן, שְׁעוֹן), ולא לה אפשר להוסיף מילים כמו קידוש ו-דיבור, שכתבתן בכתיב חסר, שהוא מקובל פחות אך אפשרי, תיצור גם כן את המבנה המורפו-אורתוגרפי 0100, ותרחיב את ההומוגרפיות, למשל קדוש (קְדוּשׁ, קְדוּשׁ), דבור (דְּבוּר, דְּבוֹר). תקצר כאן היריעה מלתאר את כלל המבנים המורפו-אורתוגרפיים המייצגים מספר תבניות מורפולוגיות, ועל כך אפשר להוסיף כי חיפוש אחר מבנים מורפו-אורתוגרפיים המייצגים תבנית מורפולוגית אחת בלבד הניב לא יותר מ-16 מבנים (Bar-On & Ravid, 2011). כדי להמחיש במקצת את היקף התופעה להלן כמה דוגמאות נוספות: אחד המבנים בעלי פוריות הומוגרפית גבוהה ביותר, ואולי אף הפורה ביותר מבין כולם הוא המבנה 000. מבנה זה מייצג חמישה מבנים תנועיים שונים ששלושה מהם מייצגים יותר מתבנית מורפולוגית אחת. אלה כוללים את משקל השם המלעילי הנפוץ מאוד *CéCeC*, הגוזר שמות מוחשיים ומופשטים (בגד, דרך, משך) וביניהן צורות רבות שחלה בהן הנמכת תנועה (מלח, פרח, קרח או נעל, צעד, זחל); את המשקל *CaCaC* הגוזר בין השאר בעלי מקצוע (זמר, ספר, כתב) ותארים (רחב, ישן, קטן), ומבנה תנועי זה תואם גם פעלים בבניין קל (ישן, קשר, ספר); את המשקל *CCaC* (כפר, כלל, כתב), שתבניתו מתאימה גם לצורת הציווי בבניין קל (ברח, שכח), ואת המשקל *CaCeC* הגוזר שמות עצם (גדר, חבר, חצר) ותארים (יפה, יבש, זקן), ותואם גם את צורת הציווי של בניין פיעל (דבר, ספר, חבר). לאלה אפשר גם להוסיף את המילים ההופכות להיות הומוגרפיות כאשר הכותב בוחר לכותבן בכתיב חסר – מילים בבניין פיעל (דבר, חבר, ספר).

מבנים מורפו-אורתוגרפיים נוספים המצטיינים בפוריות הומוגרפית הם המבנה 000, המייצג בין השאר את משקלי השם *CaCaCa* (יבשה, כתבה, מחלה), ו-*CCaCa* (ברכה, שררה, קללה), תארים בנקבה (זקנה, שמנה, מלאה), וכן פעלים בנקבה בבניין קל (כתבה, שמחה, צעקה) ובבניין פיעל (שמחה, דברה, חשבה). וכן, המבנה מ000, המייצג את המשקלים *miCCaC* (מקלט, מדבר, מספר) ו-*maCCeC* (מברג, מסמר, משדר), ואת צורת הבינוני *meCaCeC* (מדבר, מגדל, מספר), ולבסוף, המבנה הנקבי מ000, המייצג את המשקל המלעילי *miCCéCet* (משקפת, משמרת, מקטרת) ואת צורת הבינוני *meCaCéCet* (משקפת, מקטרת, מדברת).

הומוגרפיה פורה קיימת גם בתוך קטגוריית הפועל, וביטויה לרוב הוא מילים שחלופותיהן הן פעלים בבניינים שונים החולקים שורש משותף. כך למשל מילים המייצגות חלופה בצורת הסביל וחלופה בצורת הפעיל: חושב (חֹשֵׁב, חָשַׁב), עונה (עוֹנֶה/עוֹנָה, עָנָה), מונע (מוֹנֵעַ, מוֹנֵעַ); מילים המייצגות את שם הפועל בבניינים שונים, ובמקרים אלו ההומוגרפיה



נוצרת לרוב על רקע שורשים נחים או חסרים, כמו ללוות (לְלוֹת, לְלוֹת), לפנות (לְפָנוֹת, לְפָנוֹת), להראות (לְהִרְאוֹת, לְהִרְאוֹת), או מילים המייצגות פעלים תלת-עיצוריים ופעלים תנייניים, למשל, מחזר (מְחַזֵּר, מְחַזֵּר) ו-מתרגלת (מְתַרְגְּלֶת, מְתַרְגְּלֶת). צורת העתיד בפועל מזמנת אף מילים המייצגות שלוש חלופות פועליות בבניינים שונים, למשל תפתח (תִּפְתָּח, תִּפְתָּח, תִּפְתָּח), אשלה (אֲשַׁלַּח, אֲשַׁלַּח, אֲשַׁלַּח) ו-ילמד (יִלְמַד, יִלְמַד, יִלְמַד), ו-ת העתיד מזמנת הומוגרפיה עם שמות במשקלים שונים בעלי תחילית ת כמו תעלה (תַּעֲלֶה, תַּעֲלֶה), תגובה (תִּגְבֹּהַ, תִּגְבֹּהַ) ו-תמונה (תִּמְוֶנֶה, תִּמְוֶנֶה). דוגמאות אחרות בתוך מערכת הפועל הן מילים המייצגות פעלים שהם מאותו השורש ובאותו הבניין, אך נבדלים בממדי נטייה – בזמן ובמין. המקור לחלופות הנבדלות בזמן הוא האלומורפיה שנוצרת בגזירתם של פעלים משורשי נפ"ה למשל, ירד (יֵרֵד, יֵרֵד) או יצא (יֵצֵא, יֵצֵא, יֵצֵא). חלופות הנבדלות במין נוצרות משני מקורות: האחד הוא אלומורפיה בצורות הבינוני הגזורות משורשי נל"ה – שוטה (שׁוֹתֵה, שׁוֹתֵה) או חולה (חׁוֹלֵה, חׁוֹלֵה). האות הסופית ה בצורות אלה מאפשרת את פענוחן בסיומת  $a$  ליצירת הנקבה, ובסיומת  $e$  ליצירת הזכר. המקור השני הוא הסופיות  $t$  ו- $ta$  בפעלים בצורות הנקבה והזכר בגוף שני יחיד בעבר. ההומוגרפיה נוצרת כיוון ששתי הסופיות מיוצגות באופן זהה בכתב הלא מנוקד על ידי האות ת, ללא ייצוג לתנועה הסופית  $a$  בצורות הזכר: אמרת, נרדמת, דיברת, הרגשת והתעמלת. תופעה דומה חלה במבנים החבורים ההכרחיים והאופציונאליים בגוף שני יחיד ורבים: בגללך, שלך, ילדיך, אחותך ואהבתיך.

אפשר אפוא להסיק כי אופייה המשורג של המורפולוגיה העברית אחראי ליצירת מילים הומוגרפיות רבות, המייצגות חלופות הנבדלות ביניהן בתכונותיהן המורפו-תחביריות: רבות מהן מייצגות חלופות מקטגוריות לקסיקליות שונות – חלופות שמניות (ו/או תואריות) וחלופות פועליות, ובמקרים שבהם החלופות מאותה הקטגוריה, כמו קטגוריית הפועל, הן נבדלות בבניין, ולכן נבדלות בערכי הטרנזיטיביות שלהן, או בממדי הזמן והמין. אפשר להמשיך ולומר כי מעטות המילים המייצגות חלופות הנושאות תכונות מורפו-תחביריות זהות, וגם כאשר המילה מייצגת חלופות מאותה הקטגוריה הלקסיקלית הנושאות תכונות מורפו-תחביריות דומות, כמו המילים ספר (סִפֵּר, סִפֵּר), מלח (מִלַּח, מִלַּח) ו-מקלט (מִקְלָט, מִקְלָט), המייצגות כולן חלופות שמניות, הן לרוב לא מקיימות זהות מורפו-תחבירית מלאה: בשתי הדוגמאות הראשונות חלופה אחת היא שם עצם מבצע (סִפֵּר, מִלַּח) והשנייה שם עצם מוחשי, ולתכונות אלה השלכות על תפקידיהן התחביריים במשפט.

מעניין לדעת כיצד תכונות אלה משפיעות על פתרון המילה ההומוגרפית. מחקרה של בר-און (2010) בדק את השפעת הקטגוריה הלקסיקלית על פענוח מילים הומוגרפיות בודדות: 171 הנבדקים שתוארו לעיל התבקשו לקרוא רשימה של 42 מילים הומוגרפיות, שכולן ייצגו חלופה שמנית וחלופה פועלית, למשל המילים עצם (עצם, עצם), מקטרת (מקטרת, מקטרת) ו-מדבר (מדבר, מדבר). המילים סודרו ברשימה שבה כללו כל שתי מילים סמוכות חלופות במבנים תנועיים שונים, וזאת כדי למנוע השפעה של מילה אחת על זו שלאחריה, למשל גזר, מדבר, נעל, מקטרת, עצם. ממצאי המחקר העלו תמונה חד-משמעית: כל קבוצות הגיל הראו עדיפות לזיהוי החלופה השמנית על פני החלופה הפועלית. ממצאים אלה משמעותיים עוד

יותר על רקע העובדה כי המילה פוענחה כשם עצם גם במקרים שבהם החלופה הפועלית (למשל מְדַבֵּר) נשפטה בשאלוני מוכרות מקדימים כמוכרת יותר מהחלופה השמנית (מְדַבֵּר). ממצאי מחקר זה מלמדים כי הגורם המכריע בפענוח המילה הבודדת היה גורם הקטגוריה הלקסיקלית, והם תואמים מחקרים רבים על האנגלית ועל העברית העוסקים בייצוגם של שמות ופעלים במילון המנטלי, המתארים את עדיפות שם העצם על פני הפועל, הן בהקשר של רכישה הן בהקשר של משימות הפקה שונות (Childers & Berman, 2000, 2003). גורם נוסף שיכול להסביר את עדיפות שם העצם הוא ההקשר שבו בוצעה המשימה: הופעתם של שמות עצם כמילה בודדת נפוצה למדי, בעיקר כסוגי רשימות, כמו רשימת קניות, מצרכים לעוגה, מערכת שעות, רשימת חוגים ועוד. הפועל, לעומת זאת, מופיע תמיד כחלק ממשפט, ונקשר לנושא ו/או למושא, ולכן קשה יותר לזהות את המילה כמייצגת פועל כאשר היא אינה חלק ממשפט.

### מקור שני להומוגרפיה: אופייה הסינתטי של המילה העברית

הדיון עד כה עסק במילים הבסיסיות בעברית, הבנויות רובן משני צורנים: השורש והתבנית. תיאור זה נאמן באופן חלקי ביותר למציאות הטקסטואלית, הדבורה והכתובה, שבה מרבית המילים אינן מופיעות בצורתן הבסיסית, וזאת בשל אופייה הסינתטי של המורפולוגיה העברית (Berman, 1987). בעברית, כבשאר השפות השמיות, מובעות משמעויות תחביריות וצורניות רבות באמצעות צורנים המצטרפים אל המילה הבסיסית בדרך ליניארית. הביטוי הכתוב לאופי הסינתטי של המילה הוא הצטרפותן של אותיות מש"ה וכל"ב אל צידה הימני של המילה הבסיסית, והופעתן של אותיות המייצגות את צורני הגזירה הקווית, את צורני הנטייה המחויבת ואת צורני הנטייה האופציונאלית מצידה השמאלי של המילה (Ravid, 2012). כך למשל הצטרפות של האותיות ו, המייצגת צורן חיבור, ו-ב, המייצגת צורן יחס, אל צידה הימני של המילה הבסיסית זרוע, והצטרפות של האותיות ות, המייצגות צורן ריבוי, והאותיות יה, המייצגות צורן שייכות, לצידה השמאלי של המילה יוצרים את המילה הסינתטית וזרועותיה. נראה כי המילים ההומוגרפיות הנתפסות לרוב כ'טיפוסיות' לקטגוריה הן אלו שתוארו בסעיף לעיל; אם למשל נבקש ממומחים לשפה להדגים מילים הומוגרפיות, סביר להניח שהמילים שיוצעו יהיו מהדוגמאות שהובאו קודם: גזר, מדבר, ספר וכדומה. ואולם מתברר כי האופי הסינתטי של המילה העברית מייצר קבוצה מכובדת עד מאוד של מילים הומוגרפיות, שבמקרים רבים כלל אינן נתפסות ככאלה כאשר הן מופיעות כמילה בודדת.

אותיות מש"ה וכל"ב מייצגות צורני יחס (ב- כ- ל- מ-), זמן ושעבוד (ש- כש- לכש- מש-), יידוע (ה-) וקישור (ו-) (Ravid, 2005), ותפקידיהן הרבים מכתיבים את תפוצתן הרחבה בטקסט. הצטרפות של כל אחת מהאותיות למילים יוצרת פוטנציאל להומוגרפיות, וכך נוצרת קבוצה רחבה, כמו המילים בחור (בְּחֹר, בְּחֹר), בלילה (בְּלַיְלָה, בְּלַיְלָה), בזה (בְּזֶה, בְּזֶה), שמונה (שְׁמוֹנָה, שְׁמוֹנָה, שְׁמוֹנָה, שְׁמוֹנָה, שְׁמוֹנָה, שְׁמוֹנָה), שכלל (שְׁכָלֵל, שְׁכָלֵל), שקרה (שְׁקָרָה, הדבר שְׁקָרָה), שכן (שְׁכֹן, שְׁכֹן), לשבת (לְשַׁבַּת, לְשַׁבַּת), מדובר (מְדַבֵּר, מְדַבֵּר) מהדובר צה"ל

נמסר...), מפה (מִפָּה, מִפָּה), כפרי (כִּפְרִי, כִּפְרִי בשל). האות ה מייצרת כמות מרשימה במיוחד של מילים הומוגרפיות, והסיבה המרכזית לכך היא הדואליות בתפקידיה בתחילת מילה: תפקידה התחבירי כמיידעת או כמשעבדת, והיותה תחילית של תבניות מורפולוגיות רבות. כך למשל ידוע של תארים ושמות במשקל *CaCiC* (רעיל, תקיף, רגיש, רגיל, פעיל; חזיר, פקיד, צעיר, בכיר) יוצר מילים הומוגרפיות רבות שאחת מחלופותיהן היא שמות ותארים מיוחדים, והשנייה פעלים בבניין הפעיל: הרעיל, התקיף, הרגישה, הרגיל, הפעיל, החזיר, הפקידה, הצעירה. דוגמה נוספת היא ידוע של מגוון שמות ותארים המיוצגים על ידי המבנה המורפו-אורתוגרפי 000 (מתנה, שפלה, רעלה, פרשה, פסקה; שמנה, כבדה), היוצר מילים שחלופתן השנייה היא במשקל *haCCaCa*, שהוא שם הפעולה של בניין הפעיל: המתנה, השפלה, השמנה, הכבדה, הרעלה, הפסקה ו-הפרשה.

תפקידם של צורני מש"ה וכל"ב הוא לקשר בין חלקי המשפט, ולכן, מילים הכוללות צורנים אלה מופיעות תמיד כחלק מהמשפט ואין להן קיום עצמאי. נוסף על כך, הנטייה הטבעית היא לתפוס את המילה הבודדת כמייצגת יחידה מילונית אחת, כלומר יחידה שאינה מורכבת ואינה סינתטית, וזו הסיבה שקשה לזהותן כחלופה של מילה הומוגרפית. דוגמה לכך היא הצורה הנפוצה למדי המורכבת מ-ב היחס המצטרפת למילה זה. אין ספק כי צורה זו נפוצה הרבה יותר מהפועל בזה, ויחד עם זאת, כאשר נציג את המילה בזה במבודד נתקשה לראות בה את הצורה הסינתטית. לתיאור זה נמצאו גם עדויות אמפיריות: שלוש מבין 42 המילים ההומוגרפיות במחקרה של בר-און (2010), שתואר לעיל, כללו חלופה שיש בה אחת מאותיות מש"ה וכל"ב (שקר, מספר, מדבר). נוסף על כך, אחת מהמילים שניתנו לנבדקים במשימה שבה הם התבקשו להעלות את מרב החלופות למילה הייתה כפר, שאחת מחלופותיה האפשריות היא כ הדימוי המצטרפת למילה פר. איש מ-171 הנבדקים לא העלה את החלופה הכוללת את אותיות מש"ה וכל"ב בפענוח שלוש המילים הבודדות, ויותר מזה, גם כאשר התבקשו להעלות את מרב החלופות למילה כפר לא עלתה כלל האפשרות של כ הדימוי כמקדימה את המילה פר.

דוגמה נוספת למילים שאחת מחלופותיהן כמעט אינה נראית בהצגת המילה הבודדת היא הומוגרפיות הנוצרת על רקע המבנים החבורים, למשל הצורות אהבתי (אֶהֱבֹתִי), ספריה (סִפְרִיָּה), סופה (סוּפָה), ומשחקה (מִשְׁחָקֶיהָ). הצורות החבורות אינן עומדות בפני עצמן שכן הן מרפררות לישויות בטקסט, ולכן הן תמיד חלק ממנו. שתי מילים המקיימות תכונה זו – אהבתי ו-פניה – ניתנו ל-171 נבדקי המחקר והם התבקשו להעלות שתי חלופות לכל מילה. כל המבוגרים וכמעט כל תלמידי התיכון הצליחו להעלות את שתי החלופות למילים, ואולם כמעט כולם העלו בתחילה את החלופה הלא חבורה. הנבדקים הצעירים יותר התבקשו במציאת החלופה החבורה, ובין הכיתות נראתה תמונה התפתחותית, כאשר תלמידי כיתה ז' הגיעו לכ-50% הצלחה בלבד. ממצאים אלה חוזרים ומדגימים את הנטייה הטבעית לתפוס את המילה הבודדת כמייצגת יחידה מילונית אחת.

קבוצה אחרונה המודגמת כאן נוצרת על רקע הסופית ההומוגרפית ות, המייצגת את

צורן הריבוי בנקבה *ot* ואת הסיימת הנפוצה מאוד *ut*, המהווה סמן בולט לשמות מופשטים גזורי השם והתואר (Bolozky & Schwarzwald, 1992). תפוצתם הרחבה של שני הצורנים האלה יוצרת אינספור מילים הומוגרפיות, שכל אחת מהן מייצגת שם עצם או תואר בצורת הריבוי בנקבה ושם גזור מופשט. חלופותיה של המילה רגישות, לדוגמה, הן התואר בריבוי נקבה רגישות ושם עצם מופשט גזור תואר – רגישות. כך הן גם המילים פעילות, עמידות, סדירות, שכיחות, מחויבות, מקובלות, מקובעות, מופשטות, מובהקות, מוכרות, בולטות, טיפשות, חברות, כבודות וילדות, ועוד רבות אחרות, ונוסיף גם את המילה הומוגרפיות המשמשת אותנו לאורך המאמר בהקשר של מילים הומוגרפיות ותופעת ההומוגרפיות.

בנקודה זו של הדיון נחזור לקטע ההדגמה מתוך ספרו של מאיר שלו, ונסה לבחון את המילים ההומוגרפיות שבו בהתבסס על הניתוח והמיפוי שהוצע כאן. ניתוח האופי הסינתטי של המילים מראה כי יותר משליש מהמילים ההומוגרפיות בטקסט כוללות אחת או יותר מאותיות מש"ה וכל"ב, למשל, ואימתו (ואמתו), והבכירה (והבכירה), ולקלס (ולקלס), בזה (בזה) ו-החדר (החדר), ושלוש מילים הן במבנה חבור – אחי (אחי) ו-צפורני רגלי (צפורני רגלי). כאשר מילים נוספות כוללות חלופה במבנה חבור שאינה המילה בטקסט, כמו ואימתו (ואימתו), וצבעו (וצבעו). ניתוח ההבדלים בממד הקטגוריה הלקסיקלית מעלה כי החלק הארי של המילים ההומוגרפיות בטקסט (16 מתוך 22) הן כאלה הכוללות לפחות שתי חלופות מקטגוריה לקסיקלית שונה: כמחציתן כוללות חלופה שמנית או תוארית וחלופה פועלית, למשל, ואימתו, צבעה, ולקלס, הצעירה, הצבע, ישן, והיתר כוללות חלופה אחת שהיא מילת תוכן והשנייה מילת פונקציה, למשל, שיש, בזה, אבל. מעט הנותרות מייצגות חלופות מאותה הקטגוריה, והן כוללות את מילות הפונקציה את ו-אל; את התואר יפה, שאפשר לפענחו בצורת הנקבה או בצורת הזכר, ואת המילים אחי, רגלי, ציפורני. כל אחת משלוש המילים האחרונות מייצגת שלוש חלופות הדומות מאוד בתכונותיהן המורפו-תחביריות: שתי חלופות בצורת הקניין החבור, הנבדלות בממד המספר (רגלי ו-רגלי), וחלופה נוספת המשמשת כנסמך במבנה סמיכות (רגלי השולחן). ההומוגרפיה בשלושתן נוצרת על רקע חוסר העקביות בכתיבתן: יש הכותבים את צורת הרבים ב-י כפולה (רגלי) ואחרים ב-י יחידה (רגלי).

עד כה ביסס הדיון את הטענה כי מרבית המילים ההומוגרפיות מייצגות חלופות הנבדלות בערכיהן המורפו-תחביריים, וכי תכונות אלה נושאות תפקיד מכריע בפענוח המילה הבודדת. סקירת המילים מלמדת גם כי קשה לדון בהבדלי שכיחות בין חלופות המילה, כיוון שכמעט אי אפשר למצוא מילים הומוגרפיות "נקיות", כאלו ששתי חלופותיהן נושאות תכונות מורפו-תחביריות זהות, ועובדה זו מקשה על יצירת ההשוואה בין החלופות: כיצד נשווה למשל בין שתי החלופות של המילה פניה? האם נבחר להשוות בין המילה פנים למילה פניה? באופן דומה כיצד נשווה בין שתי החלופות של המילה המתנה? ועניין אחר – האופי המשורג של המילה מכתוב מילים הומוגרפיות רבות שחלופותיהן הן מאותו השורש, כלומר, מילים מאותה משפחה מורפולוגית החולקות גרעין משמעות דומה. במקרים רבים החלופות קרובות ביותר במשמעותן, והן נבדלות בעיקר בתכונותיהן המורפו-תחביריות,

כמו המילים קֶשֶׁר ו-קֶשֶׁר, מְשַׁדֵּר ו-מְשַׁדֵּר, מְשַׁחֵק ו-מְשַׁחֵק, כך גם הערכים המילוניים לאהוב ואהבה במילים אֶהְבֶּתִּיו-אֶהְבָּתִּיו, להרעיל ו-רעיל במילים הֶרְעִילָה ו-הֶרְעִילָה. תמונה זו ודאי אינה גורפת, שכן מילים מאותו גרעין משמעות יכולות להיות בדרגות שכיחות שונות, למשל שם המקום משחקה שכיח פחות משם העצם משחק, כך שיש להניח שגורם השכיחות משתתף בהשוואה בין מְשַׁחֵקָה ל-מְשַׁחֵקָה. נוסף על כך, היות ששורשים רבים נושאים כמה משמעויות, או מייצרים מילים שקשרי המשמעות ביניהן רחוקים מאוד (בר-און, 2000; גליקר, בהכנה; Berman, 2013), נוצרות חלופות מאותו שורש המייצגות ערכים מילוניים שונים, כמו לָחֵם ו-לָחֵם, פָּנִיָּה ו-פָּנִיָּה או הפֶּסְקָה ו-הפֶּסְקָה. הדוגמה האחרונה מציגה ממד נוסף המשפיע על נגישות החלופה, והוא ייצוגה הכתוב: היעדר הייצוג לתנועה *i* במילה הפסקה מבליט את החלופה הפֶּסְקָה על פני הפֶּסְקָה, וכך גם למשל במילה קבלה (קָבְלָה, קָבְלָה).

בשנים האחרונות גוברים הקולות המערערים על תפקידה הבלעדי והמרכזי של השכיחות בתהליכים של למידת המילה, ייצוגה והנגישות אליה. הטענה היא כי ההגדרה המסורתית של שכיחות כתוצר של מספר היקרויות המילה אינה מספקת, וכי הדיון בלמידת מילים ובתהלוכה צריך לכרוך יחד את ערכיה הלקסיקליים, הדקדוקיים והאורתוגרפיים, ואת הקשריה השונים (Baayen, 2010; Perfetti, 2007). הדיון בחלופות המילה ההומוגרפית מדגים טענה זו ומראה כי גורם השכיחות מאבד את מידת מרכזיותו בתחרות בין החלופות, וכי השוואה בין חלופות המילה אינה יכולה להתבונן בנפרד בתכונה כזו או אחרת, אלא היא חייבת לשכלל את צרור הגורמים הלקסיקליים, המורפו-תחביריים והאורתוגרפיים המבחינים ביניהן. תובנה זו קריטית להבנת הקריאה בצל ההומוגרפיות.

### קריאה בצל ההומוגרפיות: הבעיה היא הפתרון

השמטת הניקוד במערכת הכתב העברית יוצרת את בעיית ההומוגרפיות. הטענה המעניינת כאן היא כי הפתרון לבעיה טמון במילים עצמן: התכונות המורפו-תחביריות השונות של החלופות יוצרות סביבן הקשרים מורפו-תחביריים שונים, אשר לרוב מובילים לחלופה אחת ואינם מאפשרים את האחרת. כך למשל קורה סביב שלוש המילים ההומוגרפיות במשפט מתוך ספרו של מאיר שלו: הן קיוו שלא אבחינו בזה, אצא כך אל הבריות ואהיה ללעג ולקלס. הפועל אבחינו מצריך בסבירות גבוהה מאוד את מילת היחס ב בצירוף שם עצם או מילת פונקציה, ולכן הצורה בָּנָה היא האפשרות הצפויה. במקביל לא אפשרי להציב פועל לאחר פועל כאשר אין ביניהם צורן מקשר או פסיק; ההמשך אצא כך מכוון למילת יחס המייצגת כיוון – אָל, ולא אפשרית כאן חלופת השלילה אֶל. ולבסוף, המילה ולקלס, היושבת בתוך המבנה הרטורי אהיה ללעג ולקלס, אינה יכולה לשמש כפועל אלא רק כשם עצם. תיאור זה מעמיד במוקד ההסבר את ההקשר המורפו-תחבירי, ויש להבחין בין ההקשר המקדים את המילה וזה המופיע לאחריה: ההקשר המקדים נושא את התפקיד המרכזי בזיהוי המילה – הוא מתווה את הדרך אל הערכים הדקדוקיים והלקסיקליים של המילה המסוימת ומוביל את הקורא לפענוחה (אלון, 1995), כך שההישענות עליו אינהרנטית לתהליך הזיהוי ואוטומטית

במידה רבה. ההקשר המופיע לאחר המילה משתתף גם הוא בזיהוי המילה, אך רק זה הצמוד לה. תרומתו נובעת מהדרך שבה העיניים נעות במהלך הקריאה: תנועת העיניים אינה רציפה, אלא נעשית בדרך של קפיצות ממילה למילה, והתקבעות על כל מילה. בשלב הקיבוע (פיקסציה) העין מפיקה את המידע מהמילה, ונוסף על כך, היא מפיקה מידע לא מודע מהסימנים הסמוכים לה. נמצא כי מידע זה תורם לפענוחה של המילה הסמוכה (Deusch, Frost, Pollatsek, & Rayner, 2000; Rayner, 1998), ואפשר לשער על בסיס זאת כי הוא תורם גם לפענוח המילה עצמה.

הקשר אופטימאלי הוא הקשר מקדים המכוון לפענוחה של חלופה אחת, ואינו מאפשר כלל את החלופה האחרת. במקרה זה הדיקות ההקשר היא מוחלטת והרמטית ואין כל צורך בהקשר המופיע לאחר המילה. הקשר הדוק והרמטי יוגדר ככזה כאשר פענוח החלופה הלא נכונה ייצור הפרה דקדוקית שהאוזן (או העין) אינה יכולה לסבול, ראו למשל: 'פענוח המשפט עם החלופה הלא נכונה ייצור הפְּרָה...'. ההפרה הדקדוקית הבוטה והבולטת היא הסיבה לכך שהסיכויים לפענוח שגוי במשפטים בעלי הקשר מקדים הרמטי והדוק הם אפסיים, בפרט כאשר מדובר בקוראים בוגרים ומיומנים. מעט השגיאות שעשו הקוראים בקריאת הקטע מתוך ספרו של מאיר שלו היו אך ורק במילים שהופיעו בהקשר לא הדוק (ואלה יוצגו בהמשך), ותמונה דומה נראתה במחקרה של בר-און (2010), שבדק פענוח של מילים הומוגרפיות בתוך משפט. מכאן עולה כי כל עוד ההקשר המורפו-תחבירי המקדים מאפשר פענוח של חלופה אחת בלבד, כמעט ואין צורך במידע נוסף. הבעיה מתחילה כאשר ההקשר המקדים אינו מקיים תנאי זה, והוא מאפשר במידה כזו או אחרת חלופה נוספת. עמימות היא פתח לתחרות (Li, 2009; MacWhinney, 2001), ובהקשר של פענוח המילה ההומוגרפית תתקיים התחרות בין: (1) עצמת הכוונת של ההקשר המקדים לחלופה אחת או אחרת: ההקשר המקדים מתייחס כאן הן למידע הדקדוקי הן למידע התוכני והפרגמטי במשפט עצמו, בטקסט או מחוץ לטקסט; (2) ההקשר המופיע לאחר המילה – זה הצמוד לה וזה הרחוק יותר ו-3) תכונות המילה, דהיינו, ההבדלים בין חלופותיה. ככל שגורמים אלו יכוונו לאותה החלופה, או לפחות חלקם יכוונו לאותה החלופה והאחרים לא יסתרו אותה, כך יפחתו הסיכויים לטעויות, וככל שהם יפעלו זה כנגד זה, כלומר, יכוונו לחלופות שונות, כך תגבר התחרות בין החלופות, ויגדלו הסיכויים לטעויות.

כך בעצם נוצר רצף, שמצידו האחד משפטים בעלי הקשר הדוק והרמטי שאינם מזמנים טעויות כלל, לאורכו – משפטים המזמנים טעויות במידות שונות, בהתאם לתחרות המתקיימת בין הגורמים השונים, ובקצהו השני משפטי Garden Path. משפטים אלה מכילים דו-משמעות זמנית, הנוצרת מכך שחלקו הראשון של המשפט מכוון לתפקיד או למבנה תחבירי שאינו תואם את המשך המשפט. בדומה לתנועה בין שבילי הגן, כך הקשר המשפט מוליך שולל את הקורא ("Lead down the garden path"), ומוביל אותו ל"דרך ללא מוצא". במצב זה הקורא נאלץ לחזור לאחור ולחפש את הנתיב הנכון, כלומר לפענח מחדש את המילה בצורה המתאימה להקשר שלאחריה (Ferreira, Christianson & Hollingworth, 2001). ההומוגרפיה היא מקור פורה להיווצרותם של משפטי Garden Path בעברית, כמו למשל

המשפט האיש הולך ומדבר רחב לפניו. ההקשר המקדים את המילה מדבר מכון במידה רבה ביותר לפועל מְדַבֵּר, ורק עם המעבר לתואר רחב שלאחריו מתבררת הטעות (בר-און, 2010). הדוגמאות השונות המובאות להלן מציגות דרגות שונות של הרמטיות ההקשר. השתיים הראשונות לקוחות מהפסקה בספרו של מאיר שלו, והן מציגות הקשר שהוא הדוק דיו, אך פותח פתח לחלופה אחרת, ולכן מזמן במידה כזו או אחרת טעות פענוח (המילים ההומוגרפיות "המועדות לפורענות" מודגשות):

(1) הצחקוקים נעשו צחוקים ואימתו את חשדותי

ארבעה מבין 24 הקוראים שהשתתפו בניסוי שגו בפענוח המילה: שניים מהם החלו לקרוא אותה כשם עצם במבנה חבור ותוך כדי כך שינו מיד לפועל ("וְאִימָּ... וְאִמָּתוֹ"), ושניים אחרים פענחו את המילה כשם עצם ותיקנו מיד לפועל. דקדוקית, ההקשר המקדים מאפשר את פענוח המילה כשם עצם, כך, למשל, 'הצחקוקים נעשו צחוקים ואימתו הלכה וגברה'. פענוח זה אפשרי גם מבחינה תוכנית, שכן משפט זה פתח את קטע הקריאה, והקוראים החלו את קריאתו ללא ציפיות מוקדמות לתוכן כזה או אחר. ואולם מילת היחס את הצמודה לפועל שוללת מיידית את חלופת שם העצם, וזה ההסבר לתיקון המיידית. קשה לדעת אם חלופת הפועל זמינה יותר מחלופת שם העצם במבנה החבור, ואין ספק כי המבנה החבור אופייני ביותר לטקסט הכתוב, ובוודאי הסיפורי, כך שהוא לא מפחית מזמינות החלופה. העובדה שמרבית הנבדקים פענחו נכון את המשפט מעידה כי ההקשר המקדים כיוון במידה רבה יותר לחלופת הפועל, כאשר מילת היחס את, שהופיעה מיד לאחריו, איששה זאת באופן מלא.

(2) בעודי ישן התגנבו אלי רוני ונעמי, שתי בנותיו הקטנות של אחי, וצבעו את ציפורני רגלי ההומוגרפיה בארבע המילים המודגשות במשפט נובעת מהכתיב הלא עקבי: אפשר לכתוב את מילת היחס אלי ואת צורת הרבים של שלוש האחרות ב-י בודדת ואפשרי ב-י כפולה. פענוח נכון של כל המילים ההומוגרפיות במשפט על ידי מרבית המשתתפים בניסוי מעיד כי ההקשרים המקדימים אותן הדוקים דיים, ואולם מעט השגיאות שהופיעו מלמדות כי הם אינם הרמטיים. את פתיח המשפט אפשר לפענח כך: 'בעודי ישן התגנבו אֶלִי רוני ונעמי...'. הצורה הקולוקטיבית הספרותית התגנבו אלי מושכת למילת היחס, והיעדרו של הפסיק לאחר המילה אלי משמש מידע בר ערך עבור קוראים וכותבים אורייניים, ותומך בחלופה זו. משתתף אחד בלבד פענח את המילה כשם פרטי, ואולם, אף שהמשך המשפט – שתי בנותיו – סותר את אופן פענוחו, הוא לא תיקן את שגיאתו. השגיאה לא יצרה הפרה דקדוקית בולטת, ונראה כי זו הסיבה להיעדר התיקון. קורא אחר שגה בהמשך המשפט וקרא: 'שתי בנותיו הקטנות של אָחִי'. פענוח זה אינו סביר, שכן הצורה החבורה בנותיו מרפררת קדימה לדמות בצורת היחיד – 'בנותיו של אָחִי'. גם כאן לא תוקנה השגיאה, ונראה כי אף שהיא אינה אפשרית היא יוצרת הפרה דקדוקית 'קלה', כזו שאינה מפריעה לאוזן. שני קוראים פענחו את המילה ציפורני בצורת הקניין החבור צִפְרָנִי, ותיקנו את שגיאתם מיד עם המשך

הקריאה, כאשר הבינו כי מדובר במבנה הסמיכות ציפורני רגלי. במקרה זה התיקון התחייב, שכן בלעדיו היה נוצר רצף בלתי אפשרי של שמות עצם: 'וצבעו את צִפְרָנֵי רַגְלֵי'. ארבעה קוראים פענחו את המילה רגלי במבנה הסמיכות ציפורני רגלי בצורת היחיד צִפְרָנֵי רַגְלֵי, וייתכן כי הריבוי היחסי של השגיאות נובע מכתוב המילה, המבליט את צורת היחיד. אף שצורת היחיד צפויה פחות – מקובל יותר לדבר על ציפורני שתי הרגליים – היא אפשרית בהחלט, וזו הסיבה שאיש מהשוגים לא תיקן את טעותו. אפשר לומר כי משפט זה מדגים 'הומוגרפיות עדינה': המילים אחי, ציפורני ורגלי מייצגות חלופות שהבדלים ביניהם דקים, ולכן הן נטועות בהקשרים מורפו-תחביריים דומים, המאפשרים את שתי החלופות. פענוח שגוי במקרים אלו אינו קריטי, הוא כמעט ואינו יוצר הפרה דקדוקית ולא בהכרח פוגע בתהליך הטקסט, ולכן גם אינו גורר אחריו תיקון.

הדוגמה השלישית (מתוך בר-און, 2010) מדגימה כיצד הבדלים בולטים בין חלופות המילה עשויים לגבור על הקשר הדוק והרמטי:

### 3) המורים קבעו שהוועדה תמונה על ידי המנהלת

ההקשר המקדים כאן הוא הדוק והרמטי, שכן אין כל אפשרות למשפט: 'המורים קבעו שהוועדה תמונה...'. לכך מצטרף ההקשר הצמוד למילה – על ידי, התומך בצורת הסביל תְּמוּנָה. מה אם כך גרם לשבעה מתוך 26 תלמידים מכיתה ז' לפענח את המילה כשם עצם, ואף לשניים מבין 35 תלמידי תיכון ומבוגרים להתחיל את הפענוח במילה תְּמוּנָה, ולתקנה מיד לצורת הפועל? דוגמה זו מראה כיצד פערי חלופות בולטים באופן מיוחד יכולים לערער במקצת את ההרמטיות של ההקשר. שם העצם תְּמוּנָה, הקונקרטי, הראשוני והשכיח כל כך, "קופץ לעין" ומקדים את צורת הסביל בעתיד, הנדירה הרבה יותר. הבדלי הבולטות המשמעותיים ערערו את ההקשר המקדים בכיתה ז', ואפשר להבין זאת על רקע התפתחותו המאוחרת של מבנה הסביל (רביד, לנדאו ולובצקי, 2002). יחד עם זאת, העובדה כי מרבית התלמידים בכיתה זו, והרוב המכריע של הקוראים הבוגרים והמיומנים, לא שגו בקריאתם מלמדת כי ההקשר ההדוק היה בכל זאת הגורם המכריע בתחרות. ככל שתפחת הדיקותו כך סביר שהבדלי החלופות ישפיעו יותר, ומעניין למשל לבדוק את פענוח המשפט: קשה לדעת אם תמונה לתפקיד או לא.

דוגמאות 4-7 (מתוך בראון, בהכנה) מציגות משפטים עם הקשר מקדים ניטראלי, כלומר, הקשר המכוון במידה זהה למדי לשתי חלופות המילה. חלופות המילה גם הן לא מציגות הבדלים בולטים ביניהן, ועל רקע זה הופך ההקשר שלאחר המילה לשחקן המרכזי בפענוח:

### 4) מדי בוקר נוהגת הנערה לפנות לרופא ולהתלונן על מחושיה

### 5) מדי בוקר נוהגת הנערה לפנות את הניירת שהצטברה על השולחן

שתי חלופות המילה לפנות הן פעלים בצורת ל' שם הפועל. שני הפעלים טרנזיטיביים, וכל אחד מהם מצריך מילת יחס שונה: 'לְפָנוֹת ל/אל...', 'לְפָנוֹת את...'. מילת היחס היא חלק מידע הפועל עצמו, ובלתי נפרדת ממנו. בהקשר של מילה הומוגרפית, המייצגת פעלים



שונים, זיהוי מילת היחס הצמודה לה יוצר את ההכרעה – במקרה זה ההבחנה בין פועל בבניין קל (לִפְנוֹת) או בבניין פיעל (לְפָנוֹת). פתרון העמימות הוא מידי, מכיוון שכל שנדרש הוא זיהוי של האות התחילית של המילה הצמודה לפועל, או זיהוי מילה קצרה (את, על, אל וכדומה), והעין יכולה להסתפק בהפקת מידע מסימן אחד או שניים הסמוכים למילה.

דוגמאות 6 ו-7 מציגות משפטים דומים לאלו שלעיל, אך הפועל ומילת היחס, המאפשרת את פענוחו הנכון, מופרדים באמצעות יחידה תיאורית רחבה:

6) מדי בוקר נוהגת הנערה לפנות בהתמדה ובמהירות האפשרית לרופא, ולהתלונן על מחושיה

7) מדי בוקר נוהגת הנערה לפנות בהתמדה ובמהירות האפשרית את הניירת שהצטברה על שולחנה

בשתי הדוגמאות האלה קשה לדעת כיצד תפוענח המילה לפנות, ועדיין אין לנו מידע על כך, ואולם נוכל לומר כי במקרה של פענוח שגוי – באחד המשפטים או בשניהם – יידרש הקורא למשאבי זיכרון וקשב כדי להבחין בטעותו ולתקנה, וזאת בשל המרחק הרב שבין המילה ופתרונה.

דוגמה מספר 8 לקוחה מתוך כותרת בעיתון:

8) הנשים בנצרת הצביעו בעד הרעלה

ההקשר המורפו-תחבירי במשפט מאפשר שם פעולה ושם עצם מיועד באותה המידה, וקשה להעריך את הבדלי הבולטות בין חלופת שם הפעולה הִרְעָלָה לחלופת השם המיועד הִרְעָלָה. גם מבחינה תוכנית שתי החלופות אפשריות – אפשר לכאורה להצביע בעד הִרְעָלָה, כך שפענוח המילה מסתמך על שיקולים פרגמטיים-לוגיים בלבד, ואלה דורשים משאבי קשב והתייחסות מודעת.

אפשר להבחין כיצד הדוגמאות מתקדמות, ואולי נכון יותר מדרדרות, אל קצהו השני של הרצף, וכיצד המשפטים הופכים פחות ופחות ברי פענוח. הדוגמה האחרונה, הלקוחה גם היא מכותרת עיתון, הפכה להיות אחת "מדוגמאות הדגל" של משפטי Garden Path בעברית:

9) תמונה תובעת להטרדות מיניות בשירות המדינה

הבדלי החלופות הדרמטיים במילה תמונה הוצגו לעיל, ותוארה השפעתן האפשרית במקרים של הקשר הדוק. פענוח נכון של המשפט בדוגמה 9 הוא כנגד כל הסיכויים: מילה בראש משפט היא לרוב נושא, משמע, שם עצם, ומידע זה משתלב היטב עם חלופת שם העצם תְּמוּנָה, הבולטת כל כך. סביר להניח כי הקוראים כאן, כמו כל הקוראים הפוגשים משפט זה, הקדישו כמה שניות, ואולי יותר, בניסיון לפענח את המשפט משולל ההיגיון.

הדוגמאות שהובאו כאן מעידות כי ההקשר לא תמיד מסייע בפתרון המילה ההומוגרפית, ולעתים הוא אף מחבל בפענוחה. מידע זה מערער את אמינותו של ההקשר ומאתגר את הטענה הרווחת: קריאת העברית הלא מנוקדת מתאפשרת הודות להקשר. הדוגמאות עשויות גם לערער את הטענה שהועלתה בראשו של סעיף זה: תכונותיהן הלשוניות השונות של חלופות המילה ההומוגרפית יוצרות הקשרים מורפו-תחביריים שונים, וכך מתאפשרת הקריאה. ואולם, יש לציין כי בנייתם של משפטי Garden Path אינה פשוטה כלל וכלל (בר-און, 2010), וכך גם בנייתם של משפטים בעלי הקשר ניטראלי מקדים (בראון, בהכנה). בניית המשפטים דורשת התחכמות אל מול המבנים התחביריים המקובלים והקנוניים, וזו גם הסיבה שלרוב אנו פוגשים את משפטי ה-Garden Path בהקשר של כותרות עיתונים. כותרות עיתונאיות מנסות להעביר מידע מרבי בטקסט קצר ככל האפשר, ולכן עושות שימוש במבנים תחביריים פחות שגרתיים וצפויים (ניר, 1982), כמו המשפט שהופיע במדור רכילותי בעיתון: מהליכה להשקות לא קונים במכולת. אין ספק כי יצירתם של משפטי אד הוק לשם בדיקת תופעות, או איתור של משפטים אנקדוטליים, מעשירים את הידע וההבנה על תהליך הקריאה, אבל אינם מלמדים בהכרח על טיבו של הטקסט עצמו.

טקסט אופטימאלי יוגדר ככזה כאשר מרב המילים שבו מוקדמות על ידי הקשרים אופטימאליים, כלומר הקשרים הדוקים והרמטיים, או הדוקים דיים. הקריטריונים להדיקות ההקשר הוצגו לעיל, ובהתבסס על עקרונות אלה חזרנו לפסקה מספרו של מאיר שלו ובחנו את טיבה ההקשרי. מצאנו, כי אף לא אחד מהמשפטים הוא מסוג Garden Path. מבין 27 המילים ההומוגרפיות בטקסט (הספירה כאן היא ספירת תמנית) חמש בלבד מוקדמות על ידי הקשר שאינו הדוק והרמטי, כלומר מאפשר חלופה נוספת, אך כולן הוא הדוק דיו. אלה הן המילים שנותחו בדוגמאות 1 ו-2 לעיל. כל שאר המילים מוקדמות על ידי הקשר הרמטי והדוק. ניתוח זה תואם את המיפוי של המילים ההומוגרפיות בקטע, שהוצג בסעיף הקודם, והראה כי מרבית המילים בו מייצגות חלופות מורפו-תחביריות שונות. כמו כן, הניתוח תואם באופן מלא את הטעויות שהופיעו בניסוי: הטעויות הופיעו בחמש המילים האלה בלבד, והן היו מועטות ביותר.

מיפוי המילים ההומוגרפיות והקריטריונים להדיקות ההקשר שמוצעים כאן יכולים לשמש כלי לניתוח דרגת עמימות הטקסט. סביר להניח כי ככל שהטקסט יכיל משפטים הדוקים יותר לא ייפגע דיוק הקריאה, וככל שהוא יכיל הקשרים פחות הדוקים, כך קריאתו תהיה מורכבת יותר. במאמר הנוכחי ניתחנו קטע סיפורי אותנטי, ומצאנו כי הרוב המכריע של המילים ההומוגרפיות בו מייצגות חלופות מורפו-תחביריות שונות, וכי רובו המכריע בנוי מהקשרים הדוקים והרמטיים, ובמקומות שבהם נפתח חלון לחלופה אחרת, החלון צר מאוד. כדי לדעת אם תמונה זו אכן משקפת את הטקסט העברי יש צורך להמשיך ולהעריך טקסטים נוספים מסוגים שונים ומסוגות שונות. שאלה מעניינת נוספת, שלא נדונה במאמר הנוכחי, היא השפעתה של העמימות על קצב הקריאה. שימרון וסיון (Shimron & Sivan, 1994) מצאו שקריאה של טקסט בשפה העברית איטית יותר מקריאה של טקסט בשפה האנגלית, ואפשר אולי לייחס את ההאטה להפניית משאבי הקשב לפתרון ההומוגרפיות.

מודלים לפתרון העמימות הלקסיקלית, שהוצגו בתחילתו של המאמר, מציעים כי פענוח של חלופה דומיננטית, המוקדמת על ידי הקשר התומך בה, אינו גוזל משאבי קשב כך שתהלוכה דומה לתהלוך של מילה לא עמומה (Rayner & Frazier, 1989). באופן דומה, אפשר להניח כי קריאה של מילים הומוגרפיות בהקשרים הרמטיים כמעט ואינה דורשת הפניית משאבי קשב לפתרון העמימות, כך שקצב הקריאה אינו נפגע, וככל שההקשר פחות הדוק נדרשים משאבי קשב רבים יותר לפתרון ההומוגרפיות, ואלה ייצרו הארכה בזמן הפענוח. מחקרים עתידיים יכולים לבחון שאלה זו באמצעות בדיקת קצב הקריאה של טקסטים בעלי דרגות עמימות שונות.

לסיכום, המידע שהוצג ונאסף עד כה מלמד כי אלמלא המאפיינים הייחודיים של מרבית המילים ההומוגרפיות – היותן בעלות חלופות הנבדלות בתכונותיהן המורפו-תחביריות – לא היה יכול להיווצר סביבן הקשר התומך בפענוח. המפגש של המורפולוגיה העברית – מורפולוגיה משורגת וסינתטית, עם מערכת הכתב העברית – מערכת המייצגת את התנועות באופן עמום וחסר, אחראי ליצירת אופיין הייחודי של המילים ההומוגרפיות, ואלה מכתיבות את ההקשר המאפשר את הקריאה המוצלחת במערכת כתב עמומה. במובן זה ממחישה העברית את האמירה: "A language usually gets the sort of writing system: it deserves" (Halliday, 2003). כהמשך לכך, מעניין להמשיך ולשאל עד כמה התופעה אכן ייחודית לשפות השמיות, ועד כמה היא ייחודית למילים ההומוגרפיות? כדי לענות על שאלה זו מעניין לבחון את התכונות המורפו-תחביריות של החלופות המיוצגות במילים עמומות, כמו מילים הומופוניות או הומונימיות, וללמוד מכך על השפעת תכונות אלה על פתרון העמימות הלקסיקלית.

## ביבליוגרפיה

- אלון, ע', 1995. העברית הכתובה הבלתי מנוקדת. אוניברסיטת בן-גוריון בנגב.
- בראון, א', בהכנה. תיהלוך תחבירי בקריאה: השפעת תכונות המילה וההקשר על פענוח המילה ההומוגרפית. חיבור לשם קבלת תואר מוסמך, החוג להפרעות בתקשורת, אוניברסיטת תל-אביב.
- בר-און, ע', 2000. המציאות הפסיכולינגוויסטית של השורש אצל ילדים ומבוגרים: היבטים סמנטיים, פונולוגיים-צורניים ואורתוגרפיים. חיבור לשם קבלת תואר מוסמך, החוג להפרעות בתקשורת, אוניברסיטת תל-אביב.
- בר-און, ע', 2010. תפקיד הידע הלשוני והתפתחותו בקריאת העברית הלא מנוקדת. חיבור לשם קבלת התואר "דוקטור לפילוסופיה", בית הספר לחינוך, אוניברסיטת תל-אביב.
- בר-און, ע', 2011. תהליכים לשוניים ברכישת העברית הלא מנוקדת. בתוך ע' קורת וד' ארם (עורכות), אוריניות ושפה: יחסי גומלין, דו לשוניות וקשיים. ירושלים: מאגנס, 148-165.
- גליקר, ע', בהכנה. הטרכת שורש מודעת בפעלים עבריים: מחקר פסיכולינגוויסטי. חיבור לשם קבלת התואר "דוקטור לפילוסופיה", החוג להפרעות בתקשורת, אוניברסיטת תל-אביב.

- ניר, ר', 1982. כותרות העיתון כיחידות-שית. בתוך ש' בלום קולקה, י' טובין, ור' ניר (עורכים), עיונים בחקר השיח (עמ' 75-116). ירושלים: אקדמון, האוניברסיטה העברית.
- ניר, ר', 1993. דרכי היצירה המילונית בת זמננו. ת"א: האוניברסיטה הפתוחה.
- רביד, ד', 2010. הופעת קטגוריית הפועל בעברית בקרב פעוטות: מבט פסיכולינגוויסטי חדש. אוריינות ושפה, 3, 160-131.
- רביד, ד', 2011. כותבים מורפו(נו)לוגיה: הפסיכולינגוויסטיקה של הכתיב העברי. בתוך ע' קורת וד' ארם (עורכות), אוריינות ושפה: יחסי גומלין, דו לשוניות וקשיים. ירושלים: מאגנס, 190-210.
- רביד, ד', לנדאו, ה', ולובצקי, מ', 2002. המורפולוגיה של הסביל הפועלי בגיל בית הספר: דוגמא לרכישה מאוחרת, סקריפט, 5-6, 129-158.
- שורצולד, א', 2002. פרקים במורפולוגיה עברית. כרך ג, יחידות 7-10. האוניברסיטה הפתוחה.
- שימרון, י', 1997. תפקיד הניקוד בקריאת העברית, בתוך י' שימרון (עורך), מחקרים בפסיכולוגיה של הלשון בישראל. ירושלים: מאגנס.
- Abu-Rabia, S. (2001). The role of vowels in reading Semitic scripts: Data from Arabic and Hebrew. *Reading and Writing*, 14, 39-59.
- Baayen, R. H. (2010). Demythologizing the word frequency effect: A discriminative learning perspective. *The Mental Lexicon*, 5, 436-461.
- Bar-On, A., & Ravid, D. (2011). Morphological decoding in Hebrew pseudo words: A developmental study. *Applied Psycholinguistics*, V(32), 553-581.
- Bolozky, S., & Schwarzwald, O. R. (1992). On the derivation of Hebrew forms with the -ut suffix. *Hebrew Studies*, 33, 51-69.
- Berent, I., & Perfetti, C.A., (1995). A Rose Is a REEZ: The two-cycles model of phonology assembly in reading English. *Psychological Review*, 102 (1), 146-184.
- Berman, R.A. (1987). Productivity in the lexicon: New-word formation in Modern Hebrew, *Folia Linguistica*, 21, 225-254.
- Berman R. A. (2000). Children's innovative verbs vs. nouns: Structured elicitations and spontaneous coinages. In L. Menn & N. Bernstein-Ratner (Eds.), *Methods for Studying Language Production* (pp. 69-93). Mahwah, NJ: Lawrence Erlbaum.
- Berman R. A. (2003). Children's lexical innovations: Developmental perspectives on Hebrew verb-structure. In J. Shimron (Ed.), *Language processing and language acquisition in a root-based morphology* (pp. 243-291). Amsterdam: John Benjamins.
- Berman R. A. (2013). Revisiting roots in Hebrew: A multi-faceted view. In M. Muchnik & Z. Sadan (Eds.), *Festschrift for Ora R. Schwarzwald* (pp. 130-153). Ramat Gan: Bar-Ilan University Press.

- Breznitz, Z., & Share, D. (1992). Effects of accelerated reading rate on memory for text. *Journal of Educational Psychology*, 84, 193-199.
- Britton, B. K., 1978. Lexical ambiguity of words used in English text. *Behavior Research Methods & Instrumentation*, 10, 1-7.
- Burgess, C., & Simpson, G. B. (1988). Cerebral hemispheric mechanisms in the retrieval of ambiguous word meanings. *Brain & Language*, 33 (1), 86-103.
- Childers, J., & Tomasello, M. (2002). Two-year-olds learn novel nouns, verbs, and conventional actions from massed or distributed exposures. *Developmental Psychology*, 37, 698-705.
- Deutsch, A., Frost, R., Pollatsek, A., & Rayner, K. (2000). Early morphological effects in word recognition in Hebrew: Evidence from parafoveal preview benefit. *Language and Cognitive Processes*, 15 (4/5), 487-506.
- Ferreira, F., Christianson, K., & Hollingworth, A. (2001). Misinterpretations of Garden-Path Sentences: Implications for Models of Sentence Processing and Reanalysis. *Journal of Psycholinguistic Research*, 30(1), 3-20.
- Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin*, 123, 71-99.
- Frost, R. (2012). Towards a universal model of reading. *Behavioral and Brain sciences*, 35 (05), 263-279.
- Gentner, D. (2006). Why verbs are hard to learn. In Hirsh-Pasek, K. and Golinkoff, R. M. (Eds.), *Action Meets Word: How Children Learn Verbs* (pp. 544-564), Oxford University Press.
- Goodman, K. S. (1970). Reading: A psycholinguistic guessing game. In H. Singer & R. B. Ruddell (Eds.). *Theoretical models and processes of reading*. Newark, DE: International Reading Association.
- Halliday, M. A. K. (2003). Ideas about language. In J. Webster (Ed.), *On language and linguistics. Collected works of M. A. K. Halliday*, Vol. 3 (pp. 92-115), New York: Bloomsbury.
- Healy, A. F. (1994). Letter detection: A window to unitization and other cognitive processes in reading text. *Psychonomic Bulletin & Review*, 1(3), 333-344.
- Kambe, G., Rayner, K., & Duffy, S. A. (2001). Global context effects on processing lexically ambiguous words: Evidence from eye fixations. *Memory & Cognition*, 29 (2), 363-372.
- Li, P. (2009). Lexical Organization and Competition in First and Second Languages: Computational and Neural Mechanisms. *Cognitive Science*, 33, 629-666.

- MacWhinney, B. (2001). The Competition Model: The input, the context, and the brain. In P. Robinson (Ed.), *Cognition & Second Language Instruction*. New York: Cambridge University Press.
- Peleg, O., & Eviatar, Z. (2008). Hemispheric sensitivities to lexical and contextual information: Evidence from lexical ambiguity resolution. *Brain and Language*, 105(2), 71-82.
- Perfetti, C. A. (1985). *Reading ability*. New York: Oxford University Press.
- Perfetti, C. A. (1992). The representation problem in reading acquisition, In P. B. Gough, L. C. Ehri & R. Trieman (Eds.), *Reading acquisition* (pp. 145-174). Hillsdale, New Jersey: Lawrence Erlbaum.
- Perfetti, C. A. (2007). Reading Ability: Lexical Quality to Comprehension, *Scientific Studies of Reading*, 11(4), 1-27.
- Ravid, D. (2005). Hebrew orthography and literacy. In R. M. Joshi & P. G. Aaron (Eds.), *Handbook of orthography and literacy* (pp. 339-363). Mahwah, NJ: Lawrence Erlbaum Associates.
- Ravid, D. (2012). *Spelling morphology: The psycholinguistics of Hebrew spelling*. New York: Springer.
- Ravid, D., & Schiff, R. (2004). Learning to represent vowels in written Hebrew: Different factors across development. *First Language*, 24, 185-208.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372-422.
- Rayner, K., & Frazier, L. (1989). Selection mechanisms in reading lexically ambiguous words. *Journal of Experimental Psychology: Learning, memory and cognition*, 15, 779-790.
- Saiegh-Haddad, E. (2003). Bilingual oral reading fluency and reading comprehension: The case of Arabic/Hebrew (L1)–English (L2) readers. *Reading and Writing: An Interdisciplinary Journal*, 16, 717-736.
- Share, D. L. (2008). On the anglocentricities of current reading research and practice: The perils of overreliance on an “outlier” orthography. *Psychological Bulletin*, 134(4), 584–615.
- Shimron, J., & Sivan, T. (1994). Reading proficiency and orthography: Evidence from Hebrew and English. *Language Learning*, 44, 5-27.
- Twilley, L. C., Dixon, P., Taylor, D., & Clark, K. (1994). University of Alberta norms of relative meaning frequency for 566 homographs. *Memory & Cognition*, 22 (1), 111-126.