

דורית רביד

ביה"ס לחינוך והחוג להפרעות בתקשורת, אוניברסיטת תל אביב
ספר איריס ליון, בדפוס

כותבים מורפו(פונולוגיה): הפסיכולינגוויסטיקה של הכתיב העברי

מוקדש לאיריס ליון, מורת דרך וחברה אהובה

אחת השאלות שבלשנים נשאלים לעתים קרובות בדיון על תקשורת של לווייתנים, דולפינים או גורילות, היא "איך אתם יודעים שלבעלי חיים אין שפה"? התשובה היא פשוטה: כדי שנחשו שהתקשורת שלפנינו היא "שפה" במובן האנושי, היא צריכה להיות מורכבת מקבוצה קטנה ומוגבלת של סימנים שרירותיים ומכוונים, החוזרים על עצמם בדפוסים שיטתיים. איננו מוצאים מבנים כאלה בתקשורת בעלי חיים, אך זו כמובן הגדרה מצויינת לאורתוגרפיות (מערכות כתב) אלפביתיות, שאחת מהן היא מערכת הכתיב העברי. בפרק הזה יוצג ניתוח פסיכולינגוויסטי של ידע הכתיב העברי כידע לשוני הקיים בקרב דוברי

העברית וקוראיה משני הבטים - הבט הייצוג (representation) והבט הלמידה. התפיסה הפסיכולינגוויסטית של הלשון רואה בה מערכת ייצוג קוגניטיבי, ומתעניינת בדרך שבה ידע זה נרכש, מתארגן ומיוצג בקוגניציה של הדובר תוך אינטראקציה מתמדת עם תשומה (input) סביבתית, כמו גם בדרך שבה נעשה בידע זה שימוש הולם בהקשרים תקשורתיים שונים. הגישה לרכישת השפה שבה מעוגן פרק זה היא למידה סטטיסטית סטוקאסטית (הסתברותית): יחידות הידע נלמדות מן התשומה הלשונית של הסביבה, ומתארגנות לכלל מערכת קטגוריאלית הירארכית תוך לחץ סביבתי מתמיד, הנובע מן האופי המשתנה של הנתונים הזורמים אל התודעה בכל פרק זמן נתון. יעילות הלמידה ואיכות הייצוג של המערכת נגזרים מיחסי גומלין בין סדרה של גורמים פסיכולינגוויסטיים: שכיחות הפריטים, הקטגוריות, הדפוסים והמערכות, ושכיחות הקשרים ביניהם; שקיפותם, כלומר הקשר הברור בין צורה לבין משמעות או פונקציה; סדירות הייצוג, עקביותו והדמיון בין הפריטים (Plunkett & Bandelow, 2006). המחקר הנרחב הנשען על תפיסה זו מתמקד בעיקר בלמידת השפה הדבורה ובחקר הקריאה (Landauer & Dumas, 1997; Saffran, 2003). מטרתי העיקרית בפרק זה היא לעגן את הדיון בידע הכתיב בהקשר פסיכולינגוויסטי רחב זה, ולהצביע על הידע של כתיב העברית החדשה כעל מערכת שפתית, המעוגנת בייצוגים מורפו-פונולוגיים של העברית הדבורה מכאן, ובידע אורתוגרפי על המבנה והארגון הפנימי של מערכת הרישום של השפה העברית, מכאן. אף שהפרק הנוכחי הוא תיאורטי בעיקרו, הוא נשען על תובנות הנגזרות ממחקר אמפירי, ודיון תיאורטי זה עשוי לשפוך אור על חקר רכישת הכתיב ועל דפוסים של שגיאות כתיב בקרב ילדים ומבוגרים.

הכתיב: בן חורג לחוקרים

הקשר בין השפה הדבורה לשפה הכתובה היה תמיד נושא לדיונים ולויכוחים בין חוקרים מדיסציפלינות שונות, אך עד לאחרונה היה חקר הכתיב והתפתחותו בקרב ילדים כחלק מן הידע השפתי שלהם "בן חורג" במחקר הבלשני והפסיכולינגוויסטי. כבלשנית, התפיסה שחונכתי עליה הייתה שהלשון הדבורה היא, ורק היא, הביטוי ה"אמיתי" ליכולות השפתיות, שכן היא המאפיינת מאז ומעולם את המין האנושי,

והיא הנרכשת על ידי כל ילד¹ בחברה האנושית. לעומתה, באספקלריה היסטורית, הכתב הוא רק בן כ- 5,000 שנים, ובהבט התפתחותי, ידע השפה הכתובה אינו נחלת כל אדם בוגר. על פי תפיסה זו, הכתב אינו אלא "מראה" הרושמת את השפה הדבורה; וכך, בהיותו תופעה שניונית, אינו יכול לספק לנו תובנות חשובות על מהות השפה עצמה ועל רכישתה. כהמשך ישיר לתפיסה זו עסקו חוקרי רכישת השפה עד לעשור האחרון כמעט אך ורק בהתפתחות השפה הדבורה בקרב ילדים, ועד היום רוב רובו של המחקר הפסיכולוגי-התפתחותי מתמקד בשלוש עד חמש השנים הראשונות לחיים. אף שתחום התפתחות השפה המאוחרת - כלומר, רכישת השפה בגיל בית הספר, ובמיוחד התפתחות השיח הכתוב - הוא היום כבר תחום מחקר מבוסס (Berman, 2007), עדיין לא קל לשכנע את בלשנים להתעניין גם ברכישת הכתיב על ידי ילדים בגיל בית הספר כבחלק מן התהליך שבו הילדים רוכשים את לשונם.

בקרב פסיכולוגים ואנשי חינוך, הסיבה העיקרית לחוסר ההתעניינות בהתפתחות הכתיב כתופעה פסיכולוגית היא התפיסה של הכתיב כמייצג בעיקרו יחסים בין הפונולוגיה למערכת הכתב. מפרספקטיבה זו, הכתיב הקאנוני נתפס כמערכת המשלבת שני סוגים של יחסים בין צליל בשפה לסימן כתוב: יחסים סדירים, הכוללים שיקוף ישיר וחד-ערכי של פונמות דבורות על ידי סימני כתב - למשל, האות 'ג' מסמנת את הפונמה g ; ויחסים לא סדירים, שהם הפרה אידיאליזציה (כלומר, לא סדירה ולא שיטתית) של קשר ישיר זה, כמו למשל כתיב הפונמה f על ידי האותיות וצירופי האותיות F, PH, GH באנגלית. מכאן התפיסה של הכתיב המכיל מילים "סדירות" כמו $STAMP$, העונות על דרישת הקשר הישיר בין פונמה לגרפמה, לעומת מילים "לא סדירות" כמו $WOMEN$ באנגלית או **הוא** בעברית, שמקור הפרות הסדירות בכתיב שלהן הוא לעיתים קרובות עניין היסטורי. הסדירות מופרת על ידי יסודות הומופוניים - פונמה שאפשר לכתוב אותה בשתי גרפמות (או יותר), כמו **ת' וט'** הרושמות את t בעברית; או הדיפתונג ey שאפשר לכתוב אותו בהולנגדית כ- II או כ- EI . מנקודת ראות זו, לשם כתיב המילים הסדירות די בהכללת הקשר הגראפו-פונמי, ואילו את הכתיב הכאוטי של המילים הלא-סדירות המכילות פונמות הומופוניות צריך ללמוד בעל פה, לכל היותר בהקשר של צירופי אותיות שכיחים כמו CEI או QUI באנגלית, GLI באיטלקית. הכתיב הנכון של מילים הומופוניות, שאפשר לכתוב אותן בשתי דרכים, צריך להילמד בהקשר המילה הספציפית - למשל $red / read$ או **סתר / סטר**. לפי תפיסה זו, מעבר לעצם העניין בפריצתו של הצופן הגראפו-פונמי על ידי הילד הצעיר, למידת הכתיב עצמה תהיה כרוכה בדקדוקי עניות טרחניים של שינון כתיב מילים ודפוסי אותיות, ואין זה פלא שחקר רכישת הידע הזה לא קנה לו שוחרים רבים בין החוקרים בעלי האוריינטציה לפסיכולינגוויסטיקה התפתחותית².

מכיוון אחר לגמרי, אך עדיין בהסתמך על הקשר הגראפו-פונמי בתשתית הכתיב, גישות חינוכיות מסוימות אל האוריינות קוראות תיגר על עצם התפיסה המסורתית של הכתיב כעל מערכת קאנונית: לפי תפיסה זו, הכתיב, כמו תחומי שפה אחרים, הוא תולדה של הסכמה חברתית, ולכן אין כל יתרון מושכל לכתיב התקין של אותיות הומופוניות על פני אלטרנטיבות אחרות של כתיבה, כך שאך מקרה

¹ לאורך כל המאמר, השימוש שאני עושה במין זכר הוא גנרי, כמייצג זכר ונקבה גם יחד, ואינו מצביע על העדפות מגדריות.
² אף כי עניין רב יש לפסיכולינגוויסטים בייצוגים של מילים סדירות ולא סדירות, במסגרת המחלוקת הכללית על טיבם של מנגנוני הלמידה וההלוך של השפה.

הוא שהמילה **יתרון** כתובה בת' ולא בט'. על פי תפיסה זו, "בעלי השררה", כלומר המבוגרים האורייניים, הם אלה המכירים את הקונוונציות של הכתיב, והם אלו המכתיבים לילדים ולחסרי ההשכלה את הדרך המקובלת לכתוב אותיות הומופוניות. ואפילו להפך: לפי תפיסה זו, כשיש שתי דרכים לכתוב פונמה, הילד פועל בדרך נכונה וחכמה לא פחות מן המבוגר, כיוון שהוא בוחר באפשרות השקופה יותר, ואין למבוגר שום יתרון עליו. תפיסות אלה של הכתיב כשיקוף סדיר של הדיבור בתוספת ספיח של חריגות הנובע מהחלטות שרירותיות מסבירות את העדר העניין הכללי כלפיו בקרב חוקרי השפה. כך היה לימוד הכתיב ממודר (מלשון **מדור**) מן המחקר הבלשני והפסיכולוגיוויסטי ומודר (מלשון **הדרה**) שנים רבות אל תחום החינוך הלשוני, שמטרתו הרכשת הקונוונציות של השפה הכתובה: חוקרי חינוך ומורים פיתחו וחלקו שיטות, דרכים ופוטנטיים להקלה ולזירוז של הוראת הכתיב.

בחשיבה הפילוסופית על הלשון הדבורה והכתובה חל כבר בשנות הששים מעבר דרמטי מן הגישה שהכתב הוא תעתוק הדיבור אל האוריינות הלשונית כאוטונומית, מקור הלוגיקה והבניית החשיבה (Derwing, 1992). ואולם עד לאחרונה, חקר הכתיב לא קושר בצורה שיטתית למחקר אמפירי על ידע השפה כך שישמש כר לתובנות פסיכולינגוויסטיות חדשות על השפה ועל רכישתה. העשור האחרון הביא שינויים מרעננים בתחום זה בעולם, עם התקדמות הטכנולוגיה, הקשרים והתקשורת בין המדעים הקוגניטיביים, הבלשנות והחינוך הלשוני (Olson, 1994). מחקרים חדשים ומקיפים מתבוננים כיום בסדירויות הטמונות במערכת הכתיב מעבר לקשר פונמה-גראפמה ובחשיבות המורפולוגיה כתשתית לידע הכתיב ולהוראתו (Nunes & Bryant, 2006). בשנים אלה הוצבו, נותחו לראשונה ובוססו לולאות הקשר בין הכתיב לאפנות הדבורה (Chéreau, Gaskel & Dumay, 2007), וכן בין היכולות המורפולוגיות של ילדים ויכולות הכתיב שלהם בשפות שונות (למשל, Deacon & Bryant, 2005), כולל בשפה העברית (Levin, Ravid & Rapaport, 2001).

בהקשר הישראלי, גם תחום הלשון העברית לא נטה חסד לחקר התפתחות הכתיב כחלק מן הייצוגים הלשוניים של המשתמשים בשפה. הגישה המסורתית למחקר הלשון העברית התמקדה כמעט אך ורק בעברית הכתובה לדורותיה, ועד לעשור האחרון כמעט, לא ראתה בעברית הישראלית הדבורה או הכתובה שבפי לא-מומחים (כלומר, לא מורים, מלומדים, סופרים, עיתונאים או אנשי שם) מושא מחקר לגיטימי מעבר לדיון משעשע וקליל בתופעות סלנג. ההתייחסות לקשר בין שתי אפנויות העברית בקרב המומחים כללה בעיקרה התמקדות פרסקריפטיבית בשני סוגים של שיבושים וניסיון לתקנם: בשפה הדבורה - סטיות מן המערכת המורפו-פונולוגית ההיסטורית המיוצגת בעברית המנוקדת (למשל, המעבר בעשור האחרון מה**קָעִיל** ל-*heCCiC* כמו ב-*hezmin*); ובשפה הכתובה – בעיקר הפרות של המערכת המורכבת של כללי הכתיב המלא, המתייחסים להופעתן של אותיות אהו"י, ובמיוחד לו' ולי' (האם לכתוב **תוכנית** או **תכנית**?). בהמשך פרק זה נמצא, כי לא בכדי המיקוד הכפול הזה: קיים קשר מעניין ונסתר מן העין בין שתי התופעות שהממסד הלשוני העברי מבקש לשרשן. ואולם נקודת הראות במאמר זה היא גם מאמצי תיקון אלה הם כתופעות העשויות לספק לנו תובנה פסיכולינגוויסטית על הקשר שבין מבנה העברית לתפיסת הכתיב שלה.

הגישות הוותיקות לכתיב שתוארו לעיל, התופסות את הכתיב כמשקף את הנשמע בצורה עקבית יותר או פחות, נובעות בעיקרן מחקר האנגלית והוראתה. לעומת זאת, אנשי הלשון העברית מכירים היטב את הקשר המורכב של כתיב העברית עם רובד מופשט הרבה יותר של הדקדוק העברי – ובמיוחד הקשר עם תורת הצורות (המורפולוגיה), קשר המנוסח במפורש ברבים מכללי הכתיב של האקדמיה ללשון העברית. ואמנם יש לציין, כי מומחים ובקיאים בלשון העברית העוסקים גם בחינוך לשוני שמו להם למטרה להצביע על הקשר המושכל הזה שבין מבנה המורפולוגי של העברית לבין דרכי הוראת הכתיב (שורצולד, 1996). ואולם לעבודות חשובות אלה לא היה המשך פסיכולינגוויסטי: גשר שיטתי של מחקר היוצא מן המסורת ההיסטורית העשירה, שהיא ליבת הידע של דוברי העברית, ומקשר בין תפיסתם העכשווית את המערכת הלשונית שלהם לבין מערכת הכתיב. גם בקרב רוב הציבור של דוברי העברית, וחוקרי העברית בכלל זה, נתפס הכתיב כתחום חינוכי יבשושי קמעא, שהוא באחריותו של בית הספר: בעזרת הכתבות ועיסוק בכתיב על הילדים להגיע לכתיב הקאנוני התקין, כלומר ללא שגיאות, שכן יכולת הכתיב התקין נתפסת כמאפיין של עברית אוריינית, נחלתו של הכותב המשכיל והידוע ספר. ברצוני להראות בפרק הזה, כי יש לכתיב מה לספר לנו על הידע השפתי של דוברי העברית וכותביה גם מעבר לשיקוף מצב האוריינות של הפרט. אני מבקשת לטעון, שמערכת הכתיב היא מערכת לשונית, המבוססת על אותם העקרונות שתוארו בפתיח לפרק זה: מערכת ייצוג קוגניטיבי, הנשענת על למידה סטוקאסטית. לטענתי, מחקר על ידע הכתיב של לא-מומחים יכול ללמד אותנו על מבנה השפה, על הדרך שבה מבנים שפתיים מיוצגים בקוגניציה של דובריה, ועל התפתחות השפה הן בקרב הילדים והן מהבט היסטורי של תמורות בלשון.

ידע הכתיב כמערכת לשונית

הכול מסכימים, שהאורתוגרפיה (מערכת הכתיב) האלפאביתית היא מערכת סימבולית שבה מתקיים קשר שרירותי מוסכם בין סימני כתב לסימני דיבור. ההנחה הרגילה לגבי מערכת כזאת היא שתפקידה העיקרי ואולי היחיד הוא ליצג פונמות, וכן שהמצב הטבעי והרצוי הוא אורתוגרפיה שקופה – כלומר, שבמערכת כזאת היחס בין סימני הכתב ליחידות הפונולוגיות הוא חד-חד-ערכי: כל סימן כתב ייצג באופן ייחודי יחידה פונולוגית אחת, וכל היחידות הפונולוגיות יקבלו ייצוג בכתב. ואולם הקשר הגראפון-פונמי הוא רק אחד הקשרים שבין הדיבור לכתב. הנחת היסוד במאמר זה היא שהאורתוגרפיה אינה תמליל (transcription) של הפונטיקה בחזקת "מה שרואים הוא מה ששומעים"; כלומר, היא אינה תמונת ראי של מה שאנחנו שומעים³. ייחודה הוא בכך שהיא מערכת לשונית כתובה הפועלת בתוך המרחב הלשוני, שרובו מזוהה עם השפה הדבורה: זוהי מערכת **לשונית** מופשטת ואוטונומית בעלת קטגוריות וחוקיות פנימית משלה, הנמצאת ביחסי גומלין עם מערכות לשוניות אחרות – בעברית, במיוחד עם הפונולוגיה והמורפולוגיה. שתי תכונות מאפיינות את האורתוגרפיה העברית (כמו כל מערכת כתב אחרת) כמערכת לשונית. ראשית, כמו במערכת הפונולוגית, מספר קטן של סימני כתב – 22 אותיות ותשעה סימני ניקוד – מאפשר התארגנות של האותיות ליחידות המייצגות מילים ולהבעת אין סוף

³ עצם ההנחה כי מערכת דיבור או כתב כלשהי יכולה למסור את כל מה שיש ב"שפה" היא שגויה מיסודה (Olson, 1994).

משפטים. שנית, כמו במערכת הלשונית, מעבר לקשר השקוף שבין האות לפונמה קיים רובד ביניים נסתר (hidden units) המבטא קטגוריזציה והתארגנות לשונית ברמות עמוקות יותר – רובד המאפשר למידת כתיב מהירה ויעילה יותר מזו שהייתה מתאפשרת לו היה הקשר הגראפו-פונמי הקשר היחיד בין דיבור לכתב. רובד נסתר זה, המתארגן בעת רכישת הקריאה והכתיבה, הוא רובד הייצוג המורפולוגי, והוא גם זה המסייע ברכישת הכתיב.

אין דרך פשוטה לייצג את רבדי הידע שיש לדוברי עברית על מערכת הכתב שלהם והקשרים ביניהם. בכל ניסיון כזה ייבנה המשולש אורתוגראפיה – פונולוגיה – מורפולוגיה, היוצר קומפלקס רב ממדי, והמחייב התייחסות לאורתוגראפיה כמייצגת בעת ובעונה אחת שתי מערכות – הפונולוגיה והמורפולוגיה. הפרק הזה יעמוד על הבטים אחדים של הקשר המשולש הזה.

כדי להבין את המורכבות הזאת אפשר – ואולי רצוי – לסווג את אותיות הא"ב בשתי דרכים, כבעלות שני תפקידים שונים אך משלימים – פונולוגיים ומורפולוגיים, המתבטאים באתרים שונים במילה. ראשית נחשוף את הרובד הפונולוגי, שהוא הגלוי ביותר מכל הרבדים של הידע האורתוגרפי. מן הפרספקטיבה הפונולוגית, הכתיב העברי מייצג עיצורים ותנועות, אך לא במידה שווה: העיצורים מיוצגים במלואם ובצורה עקבית על ידי האותיות, ואילו התנועות מיוצגות בשתי דרכים – על ידי סימני הניקוד, וכן, בפחות שקיפות וסדירות, על ידי אותיות. כאן יש לציין שמערכת הכתב מתנהגת בצורה דומה לעברית הדבורה, שבה העיצורים הם יציבים ועקביים יותר מן התנועות. דוברי העברית מקבלים בשוויון נפש זוגות כמו *yérek*-ו-*yarak*, *mispara*-ו-*maspera*, *higatem*-ו-*hegatem*, *mevi*-ו-*mavi* כחלופות קבילות – למגינת לבו של הממסד הלשוני העברי. כפי שנראה בהמשך, לתנועות הפנימיות האלה ולהבדלים ביניהן אין כמעט ייצוג בכתב העברי הבלתי-מנוקד. הסבר סביר להבדל זה בעקביות, בסדירות ובבולטות של ייצוג העיצורים לעומת התנועות הן בכתב והן בדיבור הוא מרכזיותו של השורש השמי (ראו בהמשך) בעל האופי העיצורי, הנושא את ליבת המשמעות הלקסיקאלית של המילה (שורצולד, 2002; Ravid, 2003).

מבחינת הפונולוגיה, האותיות העבריות מתחלקות לשתי קבוצות עיקריות – מייצגות עיצורים ומייצגות תנועות. בקבוצה הגדולה 18 מן האותיות, המייצגות עיצורים בלבד – כמו למשל **ש'**, **ב'**, **נ'**, **ג'**, **ת'**, **ל'** או **ק'**. שגיאות כתיב באותיות אלה יהיו בעיקרן שגיאות החלפה במקרים של הומופניה – **ק' בכ'**, **ת' בט'** וכו'. אפילו ילדים צעירים בשלב הפונטי המתקדם יחליפו **פ'** ב**ב'** ולא ב**ל'** (לוי, אמסטרדמר וקורת, 1997). ניתן לומר שהמרחב הפונולוגי העיצורי יציב, מוגדר ומתוחם היטב, כלומר ברור לכותבים/קוראים מה הם העיצורים המיוצגים על ידי האותיות.

הקבוצה הקטנה של ארבע אותיות **אהו"י** היא בעלת תפקיד פונולוגי כפול: מחד גיסא, הן מייצגות סדרת עיצורים: הסותם הסדקי ², החוכך הגרוני *h*, חצי התנועה *y* והחוכך הדו-שפתי *v* – עיצורים הקרובים בתכונותיהם הפונטיות לתנועות ⁴. מאידך גיסא, אותיות אהו"י מייצגות את חמש התנועות בעברית – *a, e, i, o, u* – ובתפקידן זה הן נקראות **אימות קריאה**. לאלה תתוסף **הע'**,

⁴ העיצורים הנמוכים המסומנים בא' וה' מבוטאים בקרבה לתיבת הקול, י' מייצגת חצי תנועה, ו-ו' ייצגה היסטורית את חצי התנועה *w*.

שבמבטאה הכלל-ישראלי או לא מזרחי היא חלופה ל-א'. נקודה מרכזית בפרק זה היא שהיחסים בין אותיות אהו"י/ע ליחידות הפונולוגיות שהן מייצגות מורכבים ואינם יציבים. בתפקידן העיצורי, האותיות ו'י' נעות בין עיצור לחצי-תנועה (glide, semi-vowel). כך מסמנת האות י' הן עיצור מלא (כמו במילה ילד) והן חצי תנועה שהייצוג הפונטי שלה אינו יציב – כפי שניכר ממילים כמו ביניהם או היבט. גם הו' העיצורית תסמן הן את v והן את חצי התנועה w במילים לועזיות (דאווי, וושינגטון). לא צריך להיות חוקר כתיב כדי לשים לב ל"השתוללות" ההשמטות וההכפלות של ו'י' עיצוריות בכתיב של מבוגרים, כמו בדוגמא "איינ שום קלקול" שהופיעה בנובמבר 2007 באתר "יד שניה". א"ה(ע) בעייתיות בכיוונים אחרים: במבטא העברי הלא-מזרחי, עצם האופי העיצורי שלהן אינו בא תמיד לידי ביטוי, מותנה במבנה הרגל הפרוזודית ודורש השתדלות והדגשה מיוחדות. בדיבור השוטף שלנו מה שנותר מן הזוג CV שבו העיצור ייוצג באחת מאותיות אלה הוא לעתים קרובות התנועה בלבד – כמו למשל בהגיית המילה קראו. כבר מדיון זה באהו"י כמייצגות עיצורים ניכר, שהמרחב הפונולוגי של אותיות אהו"י עמום ואינו מוגדר היטב, ויש לכך כמובן השלכות על הייצוגים בכתיב ועל רכישתם. לנושא זה נחזור בהמשך בדיון על ממשק האורתוגרפיה-פונולוגיה-מורפולוגיה בקבוצת אהו"י כמייצגת תנועות.

הקשר אות-צליל, למרות כל מורכבותו, אינו ממצה את המידע הנמסר במערכת הכתיב העברית, ואליו מתוסף הרובד המורפולוגי המוסר משמעות, שהוא מופשט יותר. מן הפרספקטיבה המורפולוגית, אותיות הא"ב נושאות שני תפקידים עיקריים (Ravid, 2005): כל האותיות עשויות לקחת חלק בייצוג השורש השמי, ולכן ייקראו אותיות שורש – למשל, אותיות השורש גד"ר או בו"א. קבוצה אחת בתוך כלל אותיות הא"ב, הכוללת 11 אותיות בלבד (קל לזכור אותן באמצעות הנוטריקון אנ"י שלמ"ה כות"ב) מייצגות רכיבים מורפולוגיים של מוספיות תחביר, נטייה וגזירה, ולכן ייקראו אותיות פונקציה. במלה ותבואנה אלה האותיות ו'ת' וכן -נה, ובמילה כשההגדרות אלה כש-,ה',ה' וכן -ות. הרובד המורפולוגי הסמוי נחשף כאן בהתארגנות הפנימית של האותיות במילה ליחידות גדולות יותר המייצגות מורפמות (צורנים). אותיות השורש גד"ר אינן סתם רצף של שלוש אותיות בעלות אותו תפקיד – כפי שנדמה אם נישאר ברמה הפונולוגית של שלושה עיצורים: הן מצייגות יחדיו יחידה גדולה אחת המכילה את שלושתן – מורפמת השורש, שהיא יחידה לשונית בעלת משמעות, ומשמעות זו נמסרת על ידי רצף האותיות האלה, בסדר הזה, בכל הקשר שבו יופיעו. לקיומו העצמאי של צורן השורש בקוגניציה של דוברי העברית וכותביה ראיות רבות – למשל, בכך שמילים החולקות שורשים מטרימות זו את זו (Velan, Frost, Deutsch & Plaut, 2005); ביכולתנו לזהות את השורש במילה ולקשר אותו עם מילים אחרות המכילות אותו השורש גם כאשר אותיות השורש מופרדות על ידי האותיות היחידות המורשות לעשות זאת – ו'י' (ראו בהמשך), כמו במילה גידור; ביכולת לזהות את השורש ולקשר אותו למשפחה המורפולוגית שלו גם במקרים שבהם הסביבה המורפו-פונולוגית מחייבת היפוך סדר אות שורש שורקת ות' התפעל, כמו במילה הסתדר; ובכתיב הנכון של השורש גם כאשר הידמות חלקית בקוליות מסווה אחד מרכיביו, כמו במילה מסביר, העשויה להישמע בדיבור שוטף כ-mazbir (Ravid, 2003). כמו אותיות השורש, גם אותיות הפונקציה עשויות להתארגן ליחידות גדולות יותר, המייצגות צורנים שונים. כך למשל במילה ותבואנה כל אחת משתי האותיות ו'ת' היא בעלת תפקיד מורפולוגי

שונה, ומייצגת צורן נפרד: הו' היא כמובן וי"ו החיבור, והת' מסמנת בפועל את זמן העתיד יחד עם גוף ומין. ואולם אותו רצף אותיות ממש בסוף המילה **כשהגדרות** מסמן יחידה גדולה יותר - צורן הרבות - *ot*. ארבע אותיות הפונקציה הפותחות את המילה **כשהגדרות** יוצרות שלוש יחידות נפרדות - מילת הקישור **כש-**, ה' הידיעה, וה' הפותחת את משקל **הקטלה** - שם הפעולה של בניין **הפעיל**. התארגנות גמישה זו של אותיות הפונקציה מלמדת אותנו שוב עד כמה מערכת הכתב היא מערכת לשונית: מדובר במספר קטן של יסודות המסמנים מספר גדול של קטגוריות, המתקבצים יחדיו בצורה שיטתית, וניתנים לפירושים שונים בתנאים משתנים - מה שמתאפשר הודות לקיומו של רובד הביניים הנסתר בין המסמן למסומן.

לוח מס' 1 מדגים את חלוקת האותיות משתי הפרספקטיבות במילים **ותבואנה** ו-**כשהגדרות**.

<u>אותיות פונקציה המייצגות עיצורים</u>	<u>אותיות שורש המייצגות עיצורים</u>
ותבואנה כשהגדרות	ותבואנה כשהגדרות
<u>אותיות פונקציה המייצגות תנועות</u>	<u>אותיות שורש המייצגות תנועות</u>
ותבואנה כשהגדרות	ותבואנה כשהגדרות

לוח 1. תפקידי האותיות בכתב העברי בממד הפונולוגי והמורפולוגי

המבנה המורפולוגי של המילה העברית

פרק זה מבקש להבהיר עד כמה מערכת הכתיב היא מערכת שפתית אוטונומית בעלת חוקיות משלה, מערכת שהיא מחד גיסא מקְוָקָה (order) מופשטת יותר מן השפה הדבורה, ומאידך גיסא מאפשרת לנו להתבונן ישירות בכמה מן היחידות הלשוניות המרכיבות את השפה (Olson, 1994). כאשר אנו קושרים את שני מישורי המיון של האותיות לממד האורתוגרפי, כמודגם בלוח מס' 1, נראה שבאופן טיפוסי, אותיות השורש מופיעות במרכז המילה העברית הכתובה, ואילו אותיות הפונקציה יופיעו משני צידי אותיות השורש. זהו ייצוג ויזואלי של הליבה הלקסיקאלית של המילה - השורש, ושל הגיוונים הדקדוקיים המעניקים למילה את משמעותה הסופית, המופיעים במעטפת המילה. מבנה זה של המילה העברית הכתובה מוכתב על ידי אילוצי המערכות הלשוניות המובעות בה, כמו גם על ידי עקרונות אוניברסאליים של מבנה ופונקציה.

האילוצין האחד הוא המבנה המורפו-פונולוגי הפנימי של המילה העברית הגזורה. רוב המילים בעברית, כולל מילים בסיסיות השייכות ללקסיקון הליבה הנרכש בגיל צעיר ומשמש אותנו בחיי היומיום, מורכבות משני הצורנים השמיים הבסיסיים של שורש ותבנית (בניין או משקל)⁵. צורן היסוד התת-לקסיקאלי הוא השורש, הנושא את עיקר המשמעות הלקסיקאלית של המילה, וצורן התוספת המקביל הוא התבנית, המוסיפה לשורש את הקטגוריה התחבירית-סמנטית של המילה: שם עצם (**נעל, קשר**), פועל (**לימד, הקשיב**), תואר (**שבור, עגול**). התבנית מוסיפה וממיינת מילים הגזורות מאותו השורש לקטגוריות סמנטיות ספציפיות: מכשיר (**מגרדת**), פועל רפלקסיבי (**התגרד**), תואר תוצאתי (**מגורד**). השילוב בין צורן היסוד לצורן התוספת מתאפשר הודות לאופי הפונולוגי השונה שלהם. שילוב צורן

⁵ בשפות פחות עשירות מורפולוגית מן העברית, כמו האנגלית, למשל, מילים יומיומיות רבות הן חד-מורפמיות, כלומר אין להן מבנה מורפולוגי פנימי, כמודגם במילים shoe, cup, belt, teach, stop, listen.

היסוד - השורש וצורן התוספת - התבנית יוצר את המילה הבסיסית בעברית, מילים כמו **נכנס**, **פָּרַח**, **תבשיל**, **ריקוד**, **עָרַב**, **מספר** או **זָקָן**.

השורש הוא ברובו עיצורי (למשל $g-d-r$ או $l-m-d$), ואילו התבנית מכילה שתי תנועות לפחות⁶, כמו משקל קָטָל או בניין פיעל. תבניות רבות מוסיפות תחיליות ו/או סופיות למבנה התנועי הפנימי שלהן – כמו הבניינים **נפעל (נרדם)**, **הפעיל (הקשיב)**, **הופעל (הוסתר)** ו**התפעל (התרגז)**; או המשקלים **קטלן (שקרן)**, **תקטיל (תלמיד)**, **תקטלה (תרדמה)**, **מקטלת (מברשת)** או **תקטולת (תרכובת)**. תחיליות וסופיות התבנית אינן מוסיפות נפרדות כמו אלה המוכרות לנו מן האנגלית (למשל *disagreement*) – שכן הן חלק בלתי נפרד מהתבנית ובלעדיהן אין לה קיום. המבנה התנועי של המילה הבסיסית בעברית מוכתב על ידי התבנית, המסמנת היכן בדיוק ישתרגו תנועות המשקל או הבניין בין יסודות השורש – האם יוכנסו תנועות לאחר כל עיצור שורשי (כמו במקרה של **לימד**, **סיפור**, **מָלַךְ**), או שמא תתאפשר הופעתו של צרור עיצורי שורש (כמו במקרה של **משטרה**, **תלמיד**, **תכתוב**). התבנית היא גם זו הקובעת את מקום הטעם היסודי במילה העברית הגזורה שיש בה יותר מהברה אחת – האם יהיה בסוף המילה כמו ב**תלמיד** או **משטרה**, או בהברה הקודמת לסוף המילה, כמו ב**מכפלת** או **תואר**. בשפה הדבורה השורש והתבנית (בניין או משקל) משתרגים זה בזה, כך שהשורש אינו מופיע כרצף עיצורים – לעולם תפריד לפחות תנועה אחת בין שניים מעיצורי השורש (כמו ב-*talmid*).

כל מילה בסיסית בעברית – הן זו המורכבת מצורנים משורשים והן זו שאינה בעלת מבנה מורפולוגי פנימי, כמו **פרדס**, **מודם**, **עכבר** או **עץ**, ניתנת להרחבה באמצעות מוספיות קויות⁷. משמות עצם ותארים נבנות מילים ארוכות, מורכבות ולרוב גם מופשטות יותר באמצעות מוספיות גזירה כמו **-יה**, **-ות** או **-ן**. כך **טבעי מטבע**, **וטבעיות מטבעי**; **עירייה מעיר**; או **מדען ממדע**. למילים הגזורות בכל דרך ובכל קטגוריה לקסיקאלית (פעלים, תארים ושמות עצם) ניתן להוסיף צורני נטייה בעלי המשמעויות הדקדוקיות של מין, מספר וגוף. כך אנו מקבלים מילים נטויות כמו **כתבתי** ו**כתבתיו**, **מדברות**, **תעתיקי**, **טבעה**, **מחשבותיו** או **מדענינו**. סופיות הנטייה יבואו תמיד לאחר סופיות הגזירה – כלל אוניברסאלי המביע את הקשר ההדוק בין צורני הגזירה והעובדה שהם היוצרים את המילה הבסיסית, ואילו הנטייה ממשיכה ופועלת על המילה הבסיסית הזאת.

כיוון המוספיות הקויות של הגזירה ושל הנטייה הוא תמיד אחרי המילה הבסיסית, שמאלה (אם משתמשים בכתוב העברי כאמצעי המחשה) ויש לכך השלכה ברורה על מבנה המילה הכתובה. כך אנו רואים שמילה עברית תפתח באות המסמנת עיצור שורש או בעיצור תבנית (בניין או משקל), ותסתיים באות שורש או ביסוד מוספי שיכול לבוא מתבנית או ממוספית גזירה או נטייה. כך נגיע ליצירת מילים כמו **תבואנה** או **הגדרות**, אך עלינו להסביר גם איך נוצרו הדוגמאות הארוכות יותר בלוח 1 – **ותבואנה**, **כשהגדרות**. מעבר לשיקוף השורש, התבנית ומוספיות גזירה ונטייה, יקדימו את המילה הכתובה קבוצה

⁶ אני מציגה את הדברים בהכללה ובפשטות-מה, בהתייחס בעיקר לשורשים תלת-עיצוריים שלמים. המערכת המורפולוגית מורכבת בהרבה, ולמתעניינים כדאי לעיין בעבודותיהם של אורה שורצולד ושמואל בולוצקי (לדוגמא: שורצולד, 2002; Bolozky, 1999).

⁷ קוצר היריעה מונע ממני את ההתייחסות למבנים המסמנים את שינויי הזמן בפועל העברי, שהם בעיקרם משורשים.

של יסודות מורפו-תחביריים המכונים **מש"ה וכל"ב**: מילות יחס (**בכל"מ**), צורן היידוע (**ה'**)⁸, ומילות קישור (**ש-**, **ו-**), היכולות להופיע גם בצירופים כמו **כש-** או **לכש-**. בלי להיכנס לויכוח אם מדובר בצורנים או במילים, כדאי להבהיר שהקשר בין **מש"ה וכל"ב** לבין המילה שהם צמודים אליה הוא יותר מקשר כתוב: בעברית האוריינית הדבורה (המוכרת לכל תלמיד בית ספר, אם כי אינה שגורה בשימוש בקרב רוב האוכלוסיה דוברת העברית), כל אחת ואחת מהן משפיעה גם על הפונולוגיה של המילה ומושפעת ממנה, כפי שנראה בדוגמאות האלה: **בְּבִית, מְבִית, וּבֵית, לְבִית**.

לוח מס' 2 ממחיש את המבנה המורפולוגי של המילה העברית, הנובע מן התיאור דלעיל.

מש"ה וכל"ב	תחיליות תבנית	שורש	סופיות גזירה	סופיות נטייה	המילה
לכש	ת	בוא		נה	לכשתבואנה
וה	ה	גדר	(ה)	ות	וההגדרות
		טבע	י ות	ו	טבעיותו
מה	מ	פרש		ים	מהמפרשים
ב		ספר	י(ה)	ותינו	בספריותינו

לוח מס' 2. המבנה המורפולוגי של המילה העברית

הכתיב על פי אהו"י

המבנה המורפו-פונולוגי שתואר בקטע הקודם אינו משתקף בעקיבות ובצורה מלאה בשפה הכתובה (כמודגם בלוח 2) בגלל האילוץ השני המכתיב את מבנה המלה הכתובה - דפוסי הכתיב הבלתי מנוקד של העברית. בצומת הפונולוגיה, המורפולוגיה והאורתוגרפיה מוצאות חמש התנועות של העברית – a, e, i, u, o - את ייצוגן על ידי ארבע אותיות אהו"י⁹ בשתי קבוצות - **א'ה' ו-ו'י'**. פונולוגית, האותיות **א'ה'** מייצגות בעיקר את התנועות a, e כמו בדוגמאות **נפלא, שונא, מוֹנָה, מוֹנָה**¹⁰; ו' מסמנת את התנועות o, u כמו **בתורן, מובן**; ואילו י' מסמנת את i כמו **בתרכיז**¹¹. ריבוי התפקידים התנועיים של אהו"י אינו מהווה בעיה לקוראי העברית כשמדובר במילים ילידיות, מאחר שכל התנועות במילה, ולא רק אלה המובעות על ידי אהו"י, משוחזרות על סמך התבניות המורפולוגיות שבתשתית המילים, ובעזרת ההקשר התחבירי והקשר השיח. כך, למשל, הו' **במוֹבן** תיקרא כ- u בגלל תבנית הופעל בהווה. מאחר שהמבנה הפונולוגי של העברית דורש עיצור בראש המילה, **אהו"י** תנועות לא יופיעו בראש המילה. האתר האורתוגרפי שבו מופיעות **אהו"י** תנועות במידה המרובה ביותר ובגיוון התנועי הרב ביותר (כל חמש התנועות a, e, i, o, u) הוא בסוף המילה. הסיבה לכך הוא תפקידיהן המורפולוגיים של הו"י כאותיות פונקציה המובעות בתנועות (א' אינה אות פונקציה תנועית). שכיח מאוד הוא תפקידה של ה' בסימון תנועת a שבה מסתיימים משקלים רבים ממין נקבה (לדוגמא: **מקטלה** – כמו **משטרה**, **מקטלה** – כמו **מבחנה**, **תקטלה** – כמו **תבערה**, **תקטולה** – כמו **תברואה**). ואולם ה' תנועית משתתפת גם בסימון צורני גזירה אחרים, כמו סופית – **יה** (**מאפיה, שלישייה**), וצורני נטייה, כמו הכינוי החבור של

⁸ גם כאן, הדיון מוגבל ואינו כולל פונקציות אחרות של ה', כמו ה' השאלה (**הידעת?**).

⁹ הדיון בפרק זה נשען על תובנות פסיכולינגוויסטיות מן המחקר האמפירי, ואינו חופף לחלוטין את כללי האקדמיה העברית לגבי הכתיב המלא.

¹⁰ האות **א'** מופיעה לעתים בהקשרים של o (ראש) ו- u , ואולם העניין כאן הוא בדפוסים שיטתיים שהם בסיס להכללות.

¹¹ דרוש מחקר אמפירי לקבוע את מעמדה הפסיכולינגוויסטי של י' חצי תנועה המסמנת את e כמו במילה **בין**.

הנקבה (תפקידה). תרומה סטטיסטית חשובה שמורה לשורשים המסתיימים בה', והמוכרים במסורת העברית כשורשי ל"י (פנ"ה, שת"ה, בנ"ה, רא"ה וכו'). בסופם של פעלים ממין זכר הגזורים משורשים כאלה, הה' השורשית מלווה בתנועת a בזמן עבר (בכה, נענה, התרבה) וב- e בהווה ובעתיד (משווה, ישווה); פעלים, שמות ותארים ממין זכר משורשי ל"י מסתיימים לעתים קרובות בתנועת e (כמו מבנה, מעשה). גם ו"י מופיעות בשכיחות גבוהה בסוף המילה בציינן צורני מין, מספר וגוף (קחי, הפסידו, בלכתו) או את שורשי ל"י (חצי, דלי). את הא' התנועת נמצא בסוף המילה רק בתפקיד של אות שורש – כמו בדוגמאות צבא, מבוא, נביא. ריבוי תפקידים מורפולוגיים אלה וסימונם החוזר ונשנה בסופן של מילים כתובות הוא שמשאיר בנו את הרושם של שכיחות גבוהה של אהו"י בסוף המילה. מן המחקר עולה, שילדים צעירים נוטים להשמיט את סימון התנועות הסופיות באותיות אהו"י בכתיבתם (Levin & Korat, 1993; Share & Levin, 1999), אך למידת הייצוג המורפו-פונולוגי של א'ה' וו"י הוא מהיר ויחסית דל בטעויות (Ravid & Haimowitz, 2003). הופעתן הסדורה של אותיות אלה בכיתה ב' מסמנת את המעבר של הילדים לקריאה וכתיבה של העברית הבלתי מנוקדת (בר און, בהכנה). אהו"י תנועיות יופיעו בכמות פחותה בהרבה בליבת המילה, ושכיחות נמוכה זו מצטלבת עם גיוון נמוך יותר של סוגי אותיות וסוגי תנועות. כאן נמצא, שתפוצתן של א'ה' בשאר חלקי המילה הכתובה שונה מזו של ו"י: מעבר לשורש, בליבת המילה העברית יופיעו בעיקר אותיות ו"י, וייצגו בעיקר את התנועות i, o, u ולא את a, e . הסיבה לכך היא צירוף של אילוצים מורפו-פונולוגיים עם אילון אורתוגרפי: המילה העברית לא תסמן את התנועות a, e באמצעות אותיות פונקציה בליבת המילה. אבל גם הופעתן של א'ה' תנועיות כאותיות שורש בליבת המילה מוגבלת. ה' שורשית בגזרת ל"י מתחלפת בי' לפני סופית מורפולוגית (השוו בכה לבכיננו, הפנה להפנינו), וכך יוצא שאינה מופיעה באמצע המילה. הופעתה של א' שורשית תנועית בליבת המילה מוגבלת לפעלים נטויים כמו יאמר, מצאת, הפלאנו שבהם תתווסף לבסיס סופית נטייה¹². לצומת מורכב זה של קשרים וקשרי-קשרים בין המורפו-פונולוגיה לאורתוגרפיה שתי תוצאות סטטיסטיות חופפות ובעלות זיקה הדדית: האחת, שאין כמעט ייצוג לתנועות a, e בליבת המילה העברית הכתובה; והשנייה, שלא נמצא בה כמעט הופעות של א'ה'. הידע שלנו על דפוס מורכב זה יגרום לכך שכל הופעה של ה' באמצע המילה תפורש על ידינו כעיצור (למשל – בהרה, מהר, תמהיל), וכל הופעה של א' תנועית שאינה אות שורש תסמן את המילה כלועזית, כמו למשל השם סאם.

האותיות ו"י מגלות דפוס שונה מא'ה' הן מבחינה אורתוגרפית והן מבחינה מורפולוגית. להבדיל מא'ה', שיופיעו כמעט רק בסוף המילה, ו"י תנועיות מופיעות הן באמצע המילה והן בסופה. הסיבה לכך היא מורפולוגית: שתיהן משמשות הן כאותיות שורש¹³ (במילים כמו הוביל, מותר, מקום, שיר), המופיעות באמצע המילה, והן כאותיות פונקציה (הלכו, ביתו; אכלתי, בני). המופיעות בסוף המילה -- מה שמבטיח תפוצה חופשית יותר מתפוצת א'ה'. בהתבוננות מעמיקה יותר, תפוצתן של ו"י

¹² אותה א' שורשית יכולה להופיע כעיצור (יאמר, התמלאו) וכתנועה (התמלאנו) – תלוי בהקשר המורפו-פונולוגי של המילה.

¹³ אותה ו' או י' שורשית יכולה להופיע כעיצור (ויתר, ישב) וכתנועה (נותר, מושב), תלוי בהקשר המורפו-פונולוגי.

אינה חופשית באותה מידה: האות י' כמייצגת התנועה *i* מוגבלת בהופעתה בליבת המילה, והמגבלה היא מורפו-פונולוגית: היא לא תופיע לפני צרור עיצורים (במנחים של ניקוד: לפני שווא נח). זו הסיבה שהמילה *hirciz* הרגִיז מכילה שתי תנועות *i*, אך הראשונה מקדימה את הצרור *rg*, ולכן לא תיוצג בכתב, ואילו השניה, שאינה מקדימה צרור כזה, מיוצגת בכתב. ואולם הסיבה העיקרית לפרישתן הרחבה של ו'י' בכל אתרי המילה היא תפקיד מורפו-פונולוגי ייחודי שיש להן מעבר לשני התפקידים המורפולוגיים של אותיות שורש ואותיות פונקציה שתוארו לעיל. ו'י' מסמנות את **התנועות הפנימיות** *i, u, o* של תבניות – בניינים ומשקלים. בכל הדוגמאות שלהלן, ו'י' הן תנועות פנימיות כאלה: **שיקר, פוטר, הסביר, הוכנס, שודד, אכתוב, קשור, קצין, כתיב, תחלופה, תפקיד, תצרוכת, שמירה, פלוגה, עונג, ריצוף**. כפי שאפשר לראות, הן מופיעות בין אותיות השורש כפי שעושות כל התנועות בשפה הדבורה. ואולם כאן התמונה היא מונוטונית: רק ו'י' מורשות להופיע בין אותיות השורש, והן מסמנות רק את *i, u, o*. האם ניתן לייחס לו'י' פנימיות תפקיד מורפולוגי מקביל לזה של אותיות השורש ואותיות הפונקציה? נראה לי, שכאשר שוקלים את כל הרכיבים המתנים את הופעתן של התנועות הפנימיות בכתב העברי, כף המאזניים נוטה יותר לכיוון הפונולוגי מאשר המורפולוגי. ראשית, כפי שכבר ראינו, לא כל תנועות התבנית ייוצגו בכתב העברי. מבחינת כתיב הבניינים והמשקלים, *a, e* השכיחות מאוד הן תנועות נעלמות, והעדף הייצוג שלהן הוא הסיבה העיקרית לעודף העיצוריות וההומוגרפיה הרבה בעברית (זקן או זקן? גזר או גזר?) (בר און, בהכנה). במקביל, ו'י' פנימיות מופיעות במילים עבריות בעלות תבניות סדורות מורפו-פונולוגית, שקשה לקבוע לגביהן אם הן משקל בכל משמעות המושג (מה שכולל גם התייחסות למרכיב השורש): **סימן, איכר, שותף, גורל, יומן, בול, אות, פיל, סוף, חתול, תנור, סירה, חולצה, תמונה**. מכל אלה עולה, שו'י' בליבת המילה הן בעלות תפקיד פונולוגי מובהק של סימון תנועות – התפקיד הדומה ביותר לזה הקיים באורתוגרפיה האנגלית. בניגוד להן, שאר אמות הקריאה בשאר האתרים במילה העברית מוכתבים בראש ובראשונה על ידי המורפולוגיה, כלומר, הן לוקחות חלק בייצוג של צורנים. בגלל דפוסי הייצוג הברורים של המורפולוגיה בעברית, אין זה כלל בטוח שתנועות התבניות המיוצגות על ידי ו'י' (למשל, משקל **תקטיל**) אמנם נתפסות פסיכולינגוויסטית כמייצגות של חלק מתבנית¹⁴.

הבט רלוונטי המייחד את ארבע אותיות **אהו"י**, ובתוכן במיוחד את ו'י' התנועות, הוא חוסר יציבותן בכתב העברי. תיאורנו לעיל את חוסר היציבות של התנועות באפנות הדבורה של העברית, המתבטא בחילופי תבניות (משפלה / משפלה) ובשינויי תנועות (*makir / mekir*). חוסר היציבות של התנועות בכתב הוא ממין אחר, והוא מתגלה בנטייה להשמטה ולהכנסה עודפת של אותיות **אהו"י** בליבת המילה, שבשום פנים ואופן לא הייתה מתקבלת על דעת קוראי העברית וכותביה אילו היה מדובר בעיצורים. בעניין זה אני נדרשת לשני מודלים כלליים של הקשר שבין למידת הקריאה והכתיבה לבין מידע פונולוגי ואורתוגרפי. מודל אחד טוען, כי ידע הכתיב נגזר במידה רבה מיכולתם של הקוראים

¹⁴ ניתוח מעמיק יותר יגלה, ששילובי שורשים לקויים (חסרים או נחים) עם תבניות יוצרים מבנים שבהם הקורא שאינו בלשן עברי לא יוכל לדעת (וסביר להניח שגם לא יתעניין) אם האותיות האלה מייצגות יסודות שורש או תנועות פנימיות, כמו במילים **תועלת, תובנה, מושב, צורן, משימה, שיחה**.

לזהות דפוסי כתיב בעקבות חשיפות מרובות לדפוסים אלה (Katz & Frost, 2001). לפי גישה זו, יציבותו של הקשר הגראפו-פונמי היא תוצאה של קשר הדוק בין הפונולוגיה לאורתוגרפיה, קשר שאינו דורש כמות מרובה של תהלוך קוגניטיבי. לפי מודל זה, העובדה שקוראים / כותבים מוכנים לקבל סטיות מסוימות מן הכתיב הקאנוני היא עדות לכך שהייצוג המנטאלי של אותיות מסוימות הוא בלתי עקבי וחלש, ועקב כך נתפס כלא הכרחי. מודל שני (*Resonance Model*) מצביע על לולאת למידה מתמידה בין אירועים פונולוגיים לאורתוגראפיים כמנגנון האחראי לחיזוק הזיהוי של דפוסי כתיב (Stone & Van Orden, 1994). אם נוסיף את הממד המורפולוגי לקשר שבין הפונולוגיה לאורתוגרפיה בשני המודלים האלה, נוכל להסביר מדוע ייצוג אותיות אהו"י יציב יותר בסוף המילה מאשר באמצעה.

כבר ראינו שרוב ייצוג התנועות על ידי אהו"י מתנקז אל סופה של המילה העברית, וזהו הדפוס הסטטיסטי המתקבע בתודעתם של קוראי העברית וכותביה: ריבוי שיטתי זה של מגוון אהו"י סופיות מוביל ליציבות הייצוג שלהן. על יציבות הייצוג הזה מצביע לא רק הדפוס ההתפתחותי שתואר לעיל, אלא גם העובדה שהדפוס ההפוך של הכללת-יתר על ידי הוספת אם קריאה עודפת לסימון תנועת a בגוף שני, מין זכר (**אותכה, כתבתה**) שורד כשגיאת כתיב לאורך שנות בית הספר (רביד, 2006), ונושא עמו סטיגמה של חוסר אוריינות. ואולם במקביל לכך, דלות הייצוג של התנועות הפנימיות וחוסר העקביות שלו שוחקת את יציבותו הפסיכולינגוויסטית של ייצוג התנועות בליבת המילה העברית. אל הסיבות המורפו-פונולוגיות לדלות הייצוג הפנימי של תנועות יש להוסיף את כללי הכתיב המלא של האקדמיה ללשון העברית לגבי ו"י פנימיות, הנגשים בעיקר לאנשי מקצוע בתחום הלשון העברית (תשס"ב).

כללים אלה דורשים ידע פורמאלי מערכתי בניקוד ובמורפו-פונולוגיה. כך למשל הם קובעים שתנועות o , u ייוצגו על ידי ו', אך לא "במילים שבהן תנועת o מסומנת בניקוד בקמץ קטן או בחטף-קמץ שקיים בכל צורות המילה. למשל: **אמנם, אפנה, יזמה, תכנית, למחרת, צהריים** (לא כן מילים דוגמת **חֲמִירי** הנכתבת בווי"ו, משום שצורת היסוד שלה, חֲמִיר, מנוקדת בחולם). לכתובת ה' יש סייגים רבים הרבה יותר הדורשים חשיבה מופשטת ומורכבת על ניקוד מילים בגזירה ובנטייה, כמו למשל הידיעה מתי מנוקדת אות בשווא נח או מתי צורת היסוד במילה אינה בתנועת i (לֵב – לְבו).

אין בדבריי אלה שום כוונה ועניין לבקר את כללי האקדמיה לכתוב בלתי מנוקד וגם לא להציע לשנות אותם, אלא להסביר את הלולאה הסטוקאסטית המובילה לשחיקת ייצוג התנועות הפנימיות במילה. על פי שני המודלים של רכישת הכתיב וייצוגו, הקשר בין היחידה הפונולוגית והאורתוגרפית צריך להיות עקבי ורצוף בשיח הנקרא כדי לדרוש תהלוך מועט, כלומר, השגת יעילות, מהירות וחוסר מאמץ בייצוג התנועות באמצעות ו"י בכתב. ואולם התוצאה הראשונית של מימוש תקין של כללי האקדמיה בכתוב בלתי-מנוקד היא הופעה לא סדירה של אותיות אלה בתנאים שהקורא הלא-מומחה אינו יכול לאפיין ולשחזר – כלומר, בדיוק ההפך מן הדרישה להידוק לולאת הקשר בין תנועות i , o , u לבין אותיות ו"י כתנועות פנימיות בליבת המילה: אלה לא יופיעו כלל בכתוב המנוקד, ויופיעו בתנאים מורפו-פונולוגיים מורכבים בכתב הבלתי-מנוקד. התרופפות הקשר בין צורה לבין משמעות / פונקציה מוכרת לנו היטב כאחראית לתהליכי תמורות בלשון הדבורה. זה מה שקורה גם כאן: הלולאה הרופפת – כלומר, חוסר הקשר העקבי והסדיר – גורם לכותבים לחפש פתרונות עקביים משלהם. פתרון אחד, האפייני

לשנות בית הספר המוקדמות, הוא ייצוג פונולוגי מלא ככל האפשר של תנועות אלה, ובמיוחד של י' (שלהוספתה סייגים רבים יותר, כפי שתואר לעיל) – מה שמוביל לנירדם, היסביר, תיכתוב, מישטרה, תיזמורת (רביד, 2002). פתרון אחר, המאפיין כותבים בוגרים יותר, הוא השמטה עקבית של ו'י פנימיות, מה שמוביל להומוגראפיה גורפת – הרצף מספר נתפס לא רק כ-, *mispar, misper, mesaper*, *mi-séfer* אלא גם כ-*mesupar*; קשר נתפס לא רק כ-*kašar, kéšer, kašer*, אלא גם כ-*kišer, kušar* (בר און, בהכנה). השמטות אלה סובבות אותנו במילים הנכתבות בידי אורייניים יותר (קוגנטיביים) ואורייניים פחות (נקנקיות) – ויהיו בין קוראיי שלא יבחינו בבעיה במילים שבסוגריים. וכמו שקורה בתהליכים אחרים של תמורות בלשון (language change), חוסר העקביות בייצוג מתפשט מן הקבוצה המורפו-פונולוגית האחת של ו'י פנימיות שאינן בעלות תפקידים מורפולוגיים ברורים גם אל אותיות שורש, והדבר קורה בעיקר כאשר תפקידן המורפולוגי אינו ברור. כך למשל, הזמינו בתי הקולנוע את הקהל לבוא לקולנוע במודעות שקראו "פ' א' לקולנוע": כזכור, הו' בבוא היא שורשית.

בסדרת מחקרים של שיף ורביד על ייצוג האות ו' כמייצגת את התנועות *o, u* בליבת המילה (Schiff & Ravid, 2004 ואחרים) מצאנו, שתפקיד מורפולוגי ברור של האות (כמו השתתפות בצורן הרבות –ות) מבטיח את הישרדותה בכתיב המילה – כלומר, השמטה תיתפס כשגיאת כתיב. אך ו' שאינה בעלת תפקיד כזה וייצוגה משתנה בצורה לא עקבית בין הכתב המנוקד והבלתי-מנוקד, כמו הו' במשקל המלעילי קוטל (אושר, עונג) או במילים כמו שותף, סולם נתפסת בקרב הקוראים המיומנים כאופציונאלית. כלומר, הכנסתה והשמטה הופכות להיות לא-עקביות ואפילו אקראיות, ווודאי אינן נתפסות כ"שגיאת כתיב". מצאנו אפוא שקוראי העברית שווי נפש או "עיוורים" להופעתה של ו' כזאת, כלומר אינם תופסים את הופעתה או אי הופעתה כמשמעותית לכתיב. לו'י' כאלה קראנו "תנועות פנימיות אופציונאליות", ואילו לו'י' פנימיות שלא יעלה על הדעת להשמיט (כמו בתואר ברור) קראנו "תנועות פנימיות הכרחיות". הבחנה זו עשויה לסייע לחוקרים, לקלינאים ולמורים בציינון שגיאות כתיב תנועיות. לסיכום חלק זה, לוח מס' 3 מציג את מיפוי אותיות אהו"י התנועיות בשני ממדים: הממד המורפולוגי – כמייצגות אות שורש או אות פונקציה; והמדד האורתוגרפי – מיקום האות במילה ובשורש. שימו לב שהלוח אינו מציג את הפונולוגיה של אותיות אהו"י, וגם לא את תפקידיהן של אהו"י העיצוריות.

אותיות שורש תנועיות בליבת המילה	מייצגות תנועות פנימיות בין עיצורי השורש או הבסיס	אותיות פונקציה תנועיות בסוף המילה
א'ה' ו'י' גזרות פ"א (אמ"ר), ל"א (ק"א), ל"י (רא"י), ע"ו (ק"מ), ע"י (שי"ר)	ו'י'	ה' ו'י' לדוגמא: ה' נקבה (פרה), ו' זכר נסתר (בנו), י' תואר גזור שם (מעשי)
הכרחיות [לא תמיד מתממשות כתנועות בלבד; אפשר לנבא את היקרות א'ה'ו'י' כעיצור או כתנועה באותיות שורש כפונקציה של הגזרה, מיקום האות בשורש, סוגו (ישן/חדש), הבניין והנטייה]	אופציונאליות (בושם, שותף, סיבה, אפור) והכרחיות (סבוך, בהיר, דרום, מתמיד)	הכרחיות [סימון a אסור בגוף שני זכר – אמרת, קולך]

לוח 3. מיפוי אותיות אהו"י המייצגות תנועות

ייצוג צורני היסוד והתוספת בכתב

מן הקטע הקודם עולה, שהתנועות מיוצגות בעיקרן בסוף המילה הכתובה (מרָאָה, מרָאָה, מבוא, הפכי, התקדרו), ופחות בליבת המילה הכתובה. יש לכך השלכה ברורה על הייצוג השונה של השורש והתבנית בכתב. משלושת הלוחות עולה, שלצורן היסוד – השורש – יש ייצוג כתוב רציף ומלא, ולעומת זאת רק החלקים החיצוניים של צורן התוספת – התבנית – צפים ועולים ברובד האורתוגרפי של המילה. אם נתבונן במשפחה המורפולוגית הגזורה מכל אחד מארבעה שורשים – זכ"ר, דר"כ, פל"ג ולמ"ד – נוכל להדגים את השיטתיות האורתוגראפית בייצוג השורש בכתב. מן השורש זכ"ר נגזרים זכר, נזכר, הזכיר, הוזכר, זכור, זכירה, היזכרות, הזכרה, אזכרה, זיכרון, זָכַר, תזכיר, מזכר, תזכורת, מזכרת. משורש דר"כ ייגזרו דרך, נדרך, הדריך, הודרך, דריכה, הדרכה, דרוך, דריכות, דָּרַךְ, מדרכה, מדריך, מדרך (רגל), תדריך. משורש פל"ג נגזרים פילג, הפליג, התפלג, פילוג, הפלגה, התפלגות, פלוגה, פָּלַג, מפלגה, מְפַלֵּג, מְפַלֵּג, פלגנות. משורש למ"ד נגזרים למד, נלמד, לימד, התלמד, למדן, למיד, למידה, לימוד, תלמיד, תלמוד, לומדה. ההתבוננות במילים גזורות אלה מראה לנו שהשורש הכתוב מופיע ברצף במילים רבות כמו נזכר, זָכַר, מדרכה, מדרך רגל, למדן, התלמד – מלים שבהן התנועות הפנימיות הן לרוב a , e , ו- i לפני צרור עיצורים. רצף השורש הכתוב יופרע רק על ידי האותיות י' ו' המסמנות את התנועות הפנימיות i , u , o של התבנית -- כמו במקרה של זכירה, תזכיר, דרוך, מדריך, הפליג, פלוגה, לימוד, לומדה, או תלמוד.

מבנה המילה הכתובה מאיר ומחזק את צורן השורש, הנתפס כיחידה רציפה אחת. תורמים לכך גורמים אחדים. ראשית, תת-הייצוג של התנועות הפנימיות ו'י', שהן היחידות המורשות להופיע בין אותיות השורש, ממתן את האופי המשורג של השורש ומאפשר לזהות אותו בקלות כיחידה רציפה בליבת המילה הכתובה. הכתיב הלא עקבי של מילים כמו סיפור / ספור, דיבר / דבר, סולם / סלם, תזמורת / תזמורת, קובץ / קבץ מצביע על "שקיפותן" הפסיכולינגוויסטית של ו'י' פנימיות כך שאינן מפריעות לרציפות השורש. שנית, השורש הכתוב מציג חלקים שאינם באים לידי ביטוי באפנות הדבורה: לא כל השורשים הם בעלי המבנה הקאנוני התלת-עיצורי¹⁵ כמו דב"ר או רג"ל, שהתרגלנו לייחס אותו לשורש השמי. כך למשל שורשים "נחים" כמו קו"מ, בו"א, יש"ב, ות"ר מכילים יסודות שורש העשויים להתבטא כתנועות תנועיים -- מה שמביא לתת-ייצוג של השורש בדיבור. לדוגמא, המייצג היחיד של השורש בו"א במילה ותבואנה הוא העיצור v . לעומת זאת השפה הכתובה נותנת ייצוג גם ליסודות ה"נחים" או התנועיים שבשורש: תבואו, מקום, מושב, נותר, גבה. שלישית, השורש הכתוב מסמן חלופות פונולוגיות באותה אות: פ תסמן הן את p והן את f , ובמקביל, השורש הכתוב מסמן בצורה ברורה שורשים בעלי יסודות הומופוניים כמו סמך ושמה. בכל אלה תורם הייצוג הכתוב לתפיסתו של השורש כצורן שלם.

לעומת זאת, במילה הכתובה תת-ייצוג לצורן התבנית, שלעולם לא תופיע כיחידה רציפה ושלמה בכתב, ותמיד יש לשחזרה: במקרה הטוב, מחלקים נפרדים במעטפת המילה כמו נ' תחילית בנרדם, ו'.

¹⁵ ואיני מתייחסת כאן כלל לשורשים מרובעים.

סופית בשקרון, ת' ו-י' פנימית בתסביך, מ' וה' במשטרה; במקרים אחרים, אין שום סימן כתוב לתבנית: בלי ניקוד או ללא הקשר לא נוכל לדעת אם מדובר בְּהֶלֶךְ או הֶלֶךְ, זָמַר, זָמַר או זָמַר.

הייצוג הכתוב שונה אפוא מן הייצוג הדבור: ממשק האורתוגרפיה העברית עם התפקידים הפונולוגיים והמורפולוגיים של האותיות וצירופי האותיות הם שקובעים כיצד תיבנה המילה הכתובה וכיצד אנו, דוברי העברית, נייצג במחשבתנו את מבנה המילה ואת היסודות הלשוניים המיוצגים בה. ממשק זה הוא שקובע כי המילה העברית הכתובה נבנית מבפנים החוצה, ובהסתמך על דפוסי האורתוגרפיה, מציינת את רוב גבולות המורפמות בצורה ברורה יותר מאשר המילה הדבורה. בשפה הדבורה, השורש והתבנית הם צורנים משורגים; בשפה הכתובה, אופי הצורנים הוא לינארי ומובחן יותר, תוך הדרה של אותן תנועות שאינן מוצאות את ביטויין באותיות אהו". השורש מקבל ביטוי רציף וייצוג מרכזי בשפה הכתובה, וכמוהו צורנים סופיים המיוצגים באותיות פונקציה. לעומתם ייצוג התבנית התנועית בעיקרה הוא דל בכתב ודורש הפשטה וחשיבה מטאלשונית. מחקרים פסיכולינגוויסטיים התפתחותיים מספקים ראיות לייצוג יציב ובולט של השורש מגיל צעיר ולכך שייצוג התבנית מתאחר לכיתה ו' (Ravid, 2003; Ravid & Malenky, 2001; Ravid & Schiff, 2006). נראה שלייצוג הכתוב יש תפקיד חשוב בדפוסים התפתחותיים אלה.

לסיכום, ביקשתי להראות בפרק זה עד כמה נסמך ידע הכתיב בעברית על יכולות לשוניות וידע לשוני, ועד כמה רכישתו ניתנת לתיאור ולהסבר כלמידה סטטיסטית הסתברותית. לאור תיאור זה, תפקידו של הילד הלומד את הכתיב העברי הוא לאתר דפוסים קרובים ורחוקים במפה הלשונית הדבורה-כתובה ולקשר אותם זה לזה תחת לחץ בלתי פוסק של תשומה בשתי האפנויות. משימה זו מתוארת בספרות הפסיכולינגוויסטית העכשווית כ-data mining, כלומר כריית נתונים: מנגנון העושה שימוש במורכבות סביבות הלמידה הטבעיות במעקב אחר מופעים מרובים של היקרויות מזווגות של יחידות מידע כדי למצות מהן את המיפויים התשתיתיים – כלומר את ייצוגי הידע המנחים את התנהגות המערכת. מנגנונים כאלה פועלים היטב דווקא בהקשרים הנראים מורכבים וכמעט כאוטיים (Smith & Yu, 2008: 1565). נראה, כי תיאור זה הולם את התהליך של רכישת כתיב השפה העברית.

מראי מקום

- בר און, ע'. בהכנה. המעבר מכתב מנוקד לכתב בלתי מנוקד בקריאת העברית. עבודת דוקטור, אוניברסיטת תל אביב.
- האקדמיה ללשון העברית. תשס"ב. כללי הכתיב חסר הניקוד. חוברת מיוחדת של **לשוננו לעם** ותקציר באתר <http://hebrew-academy.huji.ac.il/decision4.html>
- לוין, א', פ' אמסטרדמר וע' קורת. 1997. ניצני כתיבה בגיל הגן: הבטים התפתחותיים והשפעת מבנה הלשון. בתוך י' שימרון (עורך) **מחקרים בפסיכולוגיה של הלשון בישראל: רכיש, לשון, קריאה וכתיבה**. ירושלים: מאגנס, 248-289.
- רביד, ד'. 2002. תפיסת הכתיב העברי אצל הילד בישראל: מבט התפתחותי. **מגמות**, 32, 29-57.
- רביד, ד'. 2006. כתיב עיצורים ותנועות באותיות פונקציה: מחקר התפתחותי. דו"ח מחקר מוגש ללישכת המדען הראשי, משרד החינוך, אוקטובר 2006.
- שורצולד (רודריג), א'. 1996. הוראת הכתיב בשיטת החילופים המורפופונמיים. בתוך א' שורצולד וי' שלזינגר (עורכים) **ספר הדסה קנטור: אסופת מחקרים בלשון**. רמת גן: חן, 196-204.
- שורצולד, א'. 2002. **פרקים במורפולוגיה עברית**. תל אביב: האוניברסיטה הפתוחה.
- Bolozky, S. 1999. *Measuring productivity in word formation: the case of Israeli Hebrew*. Leiden: Brill.
- Berman, R.A. 2007. Developing language knowledge and language use across adolescence. In E. Hoff & M. Shatz, eds. *Handbook of language development*. London: Blackwell, 346-367.
- Chéreau, C., M.G. Gaskel & N. Dumay. 2007. Reading spoken words: Orthographic effects in auditory priming. *Cognition*, 102, 341-360.
- Deacon, H. S. & P. Bryant. 2005. The strength of children's knowledge of the role of root morphemes in the spelling of derived words. *Journal of Child Language*, 32, 375-389.
- Derwing, B.L. 1992. Orthographic aspects of linguistic competence. In P. Downing, S. Lima & M. Noonan (Eds.), *The linguistics of literacy*. Amsterdam: John Benjamins, 193-210.
- Katz, L. & Frost, S.J. 2001. Phonology constrains the internal orthographic representation. *Reading and Writing*, 14, 297-332.
- Landauer, T. K. and Dumais, S. T. 1997. A solution to Plato's problem: the Latent Semantic Analysis theory of acquisition, induction and representation of knowledge. *Psychological Review*, 104, 211-240.
- Levin, I. & Korat, O. 1993. Sensitivity to phonological, morphological, and semantic cues in early reading and writing in Hebrew. *Merrill-Palmer Quarterly*, 39, 213-232.

- Levin, I., D. Ravid & S. Rapaport. 2001. Morphology and spelling among Hebrew-speaking children: from kindergarten to first grade. *Journal of Child language*, 28, 741-769.
- Nunes, T. & Bryant, P. 2006. *Improving literacy by teaching morphemes*. NY: Routledge.
- Olson, D. 1994. *The world on paper: The conceptual and cognitive implications of writing and reading*. Cambridge: Cambridge University Press.
- Plunkett, K. & Bandelow, S. 2006. Stochastic approaches to understanding dissociations in inflectional morphology. *Brain and Language*, 98, 194-209.
- Ravid, D. 2003. A developmental perspective on root perception in Hebrew and Palestinian Arabic. In Y. Shimron (Ed.), *Language processing and acquisition in languages of Semitic, root-based morphology*. Amsterdam: Benjamins, 293-319.
- Ravid, D. 2005. Hebrew orthography and literacy. In R.M. Joshi & P.G. Aaron (Eds.), *Handbook of orthography and literacy*. Mahwah, NJ: Erlbaum, 339-363.
- Ravid, D. & S. Haimowitz. 2006. The vowel path: Learning about vowel representation in written Hebrew. *Written Language and Literacy*, 9, 67-93.
- Ravid, D. & D. Malenky. 2001. Awareness of linear and nonlinear morphology in Hebrew: A developmental study. *First Language*, 21, 25-56.
- Ravid, D. & R. Schiff. 2006. Roots and patterns in Hebrew language development: evidence from written morphological analogies. *Reading and Writing*, 19, 789-818.
- Schiff, R. & D. Ravid. 2004. Vowel representation in written Hebrew: phonological, orthographic and morphological contexts. *Reading and Writing*, 17, 245-265.
- Share, D.L. & Levin, I. 1999. Learning to read and write in Hebrew. In *Learning to read and write*, M. Harris & G. Hatano (eds). Cambridge: Cambridge University Press.), 89-111.
- Saffran, J.R. 2003. Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science*, 12, 110-114
- Stone, G.O. & van Orden, G.C. 1994. Building a resonance framework for word recognition using design and system principles. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 1248-1268.
- Velan, H., Frost, R., Deutsch, A. & Plaut, D. 2005. The processing of root morphemes in Hebrew: Contrasting localist and distributed accounts. *Language and Cognitive Processes*, 29, 169-206.
- Yu, C. & L. B. Smith. 2007. Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science*, 18, 414-420.