

לשוננו לעם

עורך: מ' בר־אשר

מתוך

כרך מ-מא, תש"ן

קובץ לשנת הלשון



האקדמיה ללשון העברית
מוגן בזכויות יוצרים

ישראל ייבין

מפעל המילון ההיסטורי

מילון היסטורי מהו?

כידוע, הלשון משתנה במשך הזמן. יש שינויים במשמעות המלים, כגון מלת "אָבִיב", שתחילה הייתה משמעותה "תבואה לפני גמר הבשלתה", והיום משמעותה "עונת השנה שלפני הקיץ"; "נָשָׁף" הייתה תחילה "ערב", וכיום משמעותה "מסיבה". יש שינויים בצורות המלים, כגון שבלשון המקרא אומרים: אֶלֶה, ובלשון חכמים: אֵלֶּה. ה' נתכנה תחילה: הקודש ברוך הוא, ואחר כך: הקדוש ברוך הוא. יש שמלה אחת מתחלפת באחרת, כגון שתחילה אמרו "אנוכי" ואחר כך "אני"; בלשון המקרא משתמשים בכינוי הזיקה "אשר", בלשון חכמים: "ש..." וכיו"ב. מילון שבו מתועדים כל השינויים ממין זה נקרא מילון היסטורי.

אפשר להקביל מילון היסטורי לדקדוק היסטורי. מבחינים בין דקדוק סינכרוני, המתאר את הפונולוגיה והמורפולוגיה של לשון בהווה, ובין דקדוק היסטורי, המתאר את תולדות מערכות הפונולוגיה והמורפולוגיה ושינוייהן במשך הדורות. הוא הדין לגבי המילון, המתאר את אוצר המלים של הלשון. יש מילון סינכרוני, כגון זה של אברהם אבן-שושן, המתאר את הלשון המשמשת היום בדיבור ובספרות; ויש מילון דיאכרוני, כלומר היסטורי, המתאר את תולדותיה של כל מלה בלשון: את מקורה, מוצאה וגזירתה; את משמעויותיה, שימושיה וצירופיה בתקופות השונות של חיי הלשון; את צורותיה השונות, הן מצד כתיבה וניקודה הן מצד דקדוקה; ואת מידת שכיחותה בתקופות השונות. מפעלים לחיבור מילון היסטורי של הלשון הלאומית מצויים ברוב המדינות התרבותיות בעולם. יש מילונים ללשונות חיות, כגון: *Oxford English Dictionary*; ויש מילונים ללשונות שפסקו מהיות חיות בדיבור, כגון: *Middle English Dictionary*.

המילון ההיסטורי היחיד כמעט ללשון העברית המצוי בידנו הוא מילונו של אליעזר בן-יהודה. אף על פי שזה מפעל יחיד, הוא כולל כמות עצומה של עדויות מכל תקופות לשונו, הוא משתמש הרבה במקורות פרשניים

ובלשניים, מילוניים ודקדוקיים מספרותנו העתיקה, משתמש בספרות מדעית מודרנית, בכתבי-יד וקטעי גניזה מרובים, וכולל חילופי נוסח.

אכן, לפי שכבר עברו כ-80 שנה מאז התחיל מילון זה לצאת לאור, נשתנו דברים רבים: התפתח מאוד חקר תקופות רבות בלשוננו, ובייחוד חקר לשון חכמים. התפתח חקר כתבי-היד וקטעי הגניזה, ונתגלו הרבה יצירות חדשות: קטעי מדרשים, פיוטים, ועוד; נתגלו מגילות מדבר יהודה. כל זה אינו משתקף במילון בן-יהודה.

לפיכך החליטה האקדמיה ללשון העברית לפני כ-30 שנה להתחיל במפעל "המילון ההיסטורי ללשון העברית", שבו תתוארנה תולדות אוצר המלים של הלשון העברית מראשיתו ועד היום, והוא יהיה מיוסד על עיבוד ובדיקה של אלפי מקורות כתובים מראשית הלשון העברית ועד היום. נוסף על התיאור ההיסטורי, היינו: שינויי משמעויותיהן, שימושן וצורותיהן של המלים במשך ההיסטוריה, יהיו במילון גם מדור אטימולוגי, הגדרה ומציינים סמנטיים, ומדור ביבליוגרפי. כמו כן ינוצלו עדויות על צורת המלים והגייתן: תעתיקים, עדויות מנוקדות בכתבי-יד ומסורות ההגייה של עדות ישראל. "המילון ההיסטורי ללשון העברית" הוא מפעל מחקר של האקדמיה ללשון העברית. מנהלו ועורכו הוא פרופ' זאב בן-חיים, ועובדים בו כ-15 עובדים מדעיים.

יסוד עיקרי במילון היסטורי הוא קביעת מקומה של כל מובאה בסדר הזמנים. כל כרטיס מובאה מן המקורות חייב אפוא לשאת ציון כרונולוגי מדויק, עד כמה שאפשר. לפיכך נרשמות המובאות מתוך היצירות שנתחברו לפני המצאת הדפוס מתוך כתבי-יד. כל מובאה במילון מלווה שני תאריכים: תאריך חיבורה של המובאה ותאריכו של כתב-היד שלפיו הועתקה. בספרות העברית העתיקה יש לעתים הפרש גדול מאוד בין שני התאריכים, כגון: מדרש הספרא [קודם לשנת 300], כ"י וטיקן 66 לספרא [950 בקירוב]; תלמוד בבלי, מסכת זבחים [שנת 500 בקירוב], כ"י קולומביה למסכת הנ"ל [שנת 1546].

את היצירות שחוברו לאחר תקופת הדפוס מעבדים לפי דפוסן הראשון. ואפילו מאמר של אחד-העם, דרך משל, שנתפרסם בהשלח ואחר כך כונס בספרו על פרשת דרכים, מעבדים לפי הפרסום הראשון בהשלח.

היצירות שנתחברו לפני תקופת הדפוס מעובדות במפעל המילון לא לפי הדפוסים, אפילו דפוסים מקובלים כמו ש"ס וילנה, או דפוסים מדעיים, אלא לפי כתבי-היד העתיקים ביותר והמשובחים ביותר שלהן.

כידוע היום, בעיקר ממחקר לשון חכמים, בדיקת כתבי־היד חשובה מאוד לידיעת לשון מקורותינו העתיקים. נוסח הדפוסים הַעֲלִים הרבה מן התכונות המקוריות העתיקות של לשוננו על־ידי מערכת ענפה של סטנדרדיזציה והאחדה, שהשתדלה להעניק לכל שכבות לשוננו צורה אחת ותבנית אחת עד כמה שאפשר. עיבוד המקורות לפי כתבי־היד המשובחים שלהם עשוי לתת לנו תמונה נכונה יותר של המלים, משמעותן ושימושן, בכל שכבותיה של הלשון העברית. לפיכך עיבוד כל מקור ומקור לפי כתבי־היד המשובח ביותר שלו הוא עיקר גדול בעבודת "המילון ההיסטורי", ובכך אחד ההבדלים העיקריים בינו ובין שאר המילונים או הקונקורדנציות לתקופות העתיקות, שהרי הבדלים דקים בין מלים, שהם חשובים ביותר מבחינה לשונית, כגון השימוש ב"אֶת" או ב"אתה" לכינוי הנוכח; חילופי בתחילה – כתחילה; הללו – האילו; מתי – אימתי; הימנו – ממנו; התפעל – נתפעל, ניטשטשו עד כדי כך בדפוסים וכן במילונים ובקונקורדנציות הערוכים לפיהם, עד שאינם מאפשרים כלל בדיקה לשונית. ידיעת מעמדן ושימושן של כל התיבות הללו אפשרית רק לפי כתבי־יד.

עיקר גדול אחר בעבודת "המילון ההיסטורי" הוא השימוש במחשב. כבר בראשית העבודה, בשעה שהשימוש במחשב עדיין לא היה מקובל בעבודות מילוניות בעולם, נתייסדה העבודה על השימוש במחשב. דבר זה יש לזקוף לזכותו של פרופ' בן־חיים, מגשים רעיון המילון. היום השימוש במחשב כבר מקובל ברוב המפעלים המילוניים הגדולים בעולם. ארבעה יתרונות לעבודה במחשב:

- א. המחשב מסייע סיוע רב בעיבוד החומר. עריכת קונקורדנציה של טקסט בעזרת מחשב נמשכת חלק זעיר מן הזמן הדרוש לעשייתה בעבודת ידיים.
- ב. רק במחשב אפשר להחסין כמויות כה עצומות של חומר: מיליוני תיבות על הקשרן, ניתוחן, מראה־מקומן, וכו', כדרוש לחיבור מילון היסטורי.
- ג. המחשב מאפשר עדכון שוטף ומתמיד של החומר שבמפעל. נסתיים איסופו של חומר לתקופה מסוימת, ואחר כך נתגלתה יצירה נוספת שיש לשלבה באותה תקופה, ויש צורך לעבדה ולרשום את מליה בתוך החומר שעבודו נסתיים – המחשב קולט יצירה זו ומשלב את החומר שבה במאגר הקיים בלא כל קושי. כמו כן כל תיקון, כל שינוי בהבנת המלה, בקריאתה ובניתוחה הדקדוקי, שיש צורך להוסיף למאגר הקיים, אפשר לעשות במחשב במהירות ובלא כל פעולה מיוחדת. החומר שעובד ונאגר במחשב מעודכן תמיד.

ד. יתרון חשוב ביותר: המחשב מאפשר עיבודים מעיבודים שונים של החומר הנאגר בו. דרך משל: עיבוד דקדוקי. המחשב עשוי לספק את כל חומר הצורות שבמאגרו בצורת פרדיגמה של הפועל לנטיותיו, השם לנטיותיו, וכו'. הוא עשוי לספק רשימות לקסיקולוגיות, כגון שכיחות מלים, השוואת שכיחות מלים בתקופה אחת לשכיחותן בחברתה. רשימות של הצטרפויות שונות של מלים ובדיקת שכיחותן כדי לגלות קולוקציות, ניבים, מלים מוצרכות על-ידי הפועל וכיו"ב.

לעת עתה מפעל המילון ההיסטורי עוסק בעיבודן של שתי תקופות בלבד מתקופות לשוננו: התקופה העתיקה, מחתימת המקרא ועד מוצאי תקופת הגאונים (1000 בקירוב), והתקופה החדשה, משנת 1750 ועד היום. מחוסר אמצעים עדיין לא הוחל בעיבוד תקופת המקרא והתקופות שבין 1000 ל-1750. לשתי התקופות הנזכרות פרסם המפעל ספרי מקורות. לתקופה העתיקה: **ספר המקורות, חלק ראשון: למן חתימת המקרא ועד מוצאי תקופת הגאונים** (ירושלים תשכ"ג, תשל"ז), ובו רשימת כל החיבורים שנתחברו בתקופה העתיקה, כתבי-היד שלהם ומהדורות מדעיות; לתקופה החדשה: **ספר המקורות לתקופה שמן 1750 ואילך**. מחברת ראשונה: מבחר חיבורים מן הספרות היפה בעברית (1860–1920), ירושלים תשל"ז.

היצירות מן התקופות הקדומות הן אך שריד של חיי לשון שלא נודעו לנו, ויש להן מעמד מיוחד בלשוננו בהשפעה על עיצוב הלשון שלאחר זמןן ואף בימינו. לפיכך הן ראויות לעיבוד בשיטת המיצוי המלא, היינו: התקנת קונקורדנציות מלאות לפי כתבי-יד משובחים. עיבוד בשיטת קונקורדנציה הוא אובייקטיבי ואינו מושפע משיקול דעתו של העובד, איזו מלה לרשום ואיזו מלה לא לרשום. כמו כן עיבוד בשיטת קונקורדנציה, כשהוא נעשה במחשב, נעשה פשוט, קל ומהיר יותר משהיה. לפיכך הוחלט במערכת המפעל לעבד כמעט את כל יצירות התקופה העתיקה בשיטת קונקורדנציה מלאה.

שפע היצירות שמתקופות מאוחרות מונע נקיטת דרך המיצוי השלם. בתקופות אלו אי אפשר לכלול בעיבוד לצורך המילון אלא חלק מסוים מן החומר שבמקורות, שאם לא כן לא תסתיים הכנת החומר למילון בזמן מתקבל על הדעת. כנגד זה, בעיית הברירה, האיסוף שבמבחר, איזה מקור לקרב ואיזה לרחק, מסובכת וטעונה שיקול רב על-פי עקרונות קפדניים. חומר המקורות המעובד בשיטת קונקורדנציה בספרות העתיקה עתיד לכלול כ-5,500,000 תיבות, בסוגי ספרות אלו:

50,000	מגילות מדבר יהודה וכתובות
3,400,000	ספרות התנאים, תלמוד ומדרש
550,000	גאונים
550,000	קראים
850,000	פיוט
100,000	מסורה ודקדוק
5,500,000	

לגבי רובו של החומר כבר נסתיימו כל שלבי העיבוד והוא מוכן לקונקורדנציה. יש לקוות שבתוך 3–4 שנים יסתיים עיבודו של כל החומר. בספרות החדשה מעובדות יצירות שבהן למעלה מ-4,000,000 תיבות, בסוגי ספרות אלו: הספרות היפה לסוגיה; פובליציסטיקה (מאמרים, מסות, פליטונים וקטעי עיתונות); ספרות המוסר וחסידות; מדע-הטבע ורפואה; היסטוריה, גאוגרפיה ורשמי מסעות; ביוגרפיה, איגרות ואגרונים. חיבורים הכוללים 1,600,000 תיבות כבר נסתיים עיבודם לכדי קונקורדנציה, והשאר נתונים בתהליכי העיבוד.

עד היום העבודה ב"מפעל המילון" מתרכזת באיסוף חומר העדויות למילון, ועדיין לא הוחל בעיבודן למילון עצמו. דוגמה אחת של המילון יצאה לאור בשנת תשמ"ב (בלשונו, כך מו, חוברת ג-ד), בעריכת פרופ' זאב בן-חיים, ובה השורש ער"ב על כל הערכים הנגזרים ממנו.

בינתיים מצטבר במפעל "המילון ההיסטורי" חומר עדויות בכמות עצומה. חומר זה נאסף בצורה אובייקטיבית, על ידי עיבוד מלא של היצירות בלא התערבות סובייקטיבית כלשהי; הוא ערוך בשיטה מדעית קפדנית ביותר: קריאה מדויקת של נוסח כתב-היד ומסירתו כדיוקן, ניתוח המלים לפי הפרשנות והספרות המדעית המודרנית ביותר; אפשרות עדכון שוטף ומתמיד. עצם האוסף הזה חשוב ביותר, והיקפו רחב יותר מזה שייכלל במילון ההיסטורי. הוא נמצא על סרטי מחשב, ולפיכך ניתן לעיבודים שונים. היינו: זה מעין ארכיון, הכולל כמות עצומה של חומר מדויק ומבוסס לתולדות הלשון העברית. מתוכו יכתב המילון ההיסטורי, אבל מתוכו אפשר גם לערוך מחקרים ועבודות שונות אחרות, כגון בדיקת תופעות דקדוקיות, סמנטיות וכיו"ב; תיאור לשונו של מקור מסוים אחד; לשונה של תקופה אחת, וכיו"ב. זהו ארכיון קבוע, מפעל קבוע שבו הולך ומצטבר חומר רב, וכל הזמן הוא אקטואלי וניתן לשימוש לחוקרים ולמעיינים. המילון ההיסטורי הוא עדיין תכלית עבודת המפעל וגולת-

הכותרת שלה, אך הוא חזון לשנים הבאות. ואילו ארכיון-הקבע של הלשון מתחיל לקרום עור וגידים, והוא כבר קיים.

מתוך הארכיון הזה הוציאה לאור זה לא מכבר מערכת המילון ההיסטורי את חומרים למילון – סדרה א' – קונקורדנציה מלאה, ערוכה לפי שורשים, של כל המקורות שנתחברו בעברית מאחרי תקופת המקרא ועד סוף תקופת התנאים. הקונקורדנציה מבוססת על כ-820,000 תיבות. נוסף על קונקורדנציה נתפרסמו גם הטקסטים של המקורות, בדוקים לפי כתבי-היד המשובחים ביותר שלהם, וכן רשימות לקסיקולוגיות: המלים לפי שכיחותן, הפעלים לבנייניהם, השמות למשקליהם, וכיו"ב.

דוגמאות לתולדות המלים המשתקפות בארכיון שב"מפעל המילון" (ועד סוף תקופת התנאים):

תיבת "מָעַט" משמשת במקרא כשם-עצם, כגון: "מעט מים" (בראשית כד, יז), כתואר פועל, כגון: "אחאב עבד את הבעל מעט" (מלכים ב' י, יח), ופעמיים כתואר שם, כגון: "יהיו ימיו מעטים" (תהלים פט, ח). שימוש כזה עדיין מצוי במגילות מדבר יהודה. אך בלשון התנאים "מעט" נדיר, ומשמש רק כתואר-פועל (6 פעמים בקונקורדנציה). ואילו כתואר שם משמשים: "מעוט", "מעוטה", "מעוטים" (51 פעמים, כגון: "שנופו מרובה ושרשיו מעוטינ", אבות ג, יז); "ממועט", "ממועטים" וכו' (50 פעם, כגון: "אחד מרובים ואחד ממועטים", ברכות ז, א); וכן "מועט" (79 פעמים, כגון: "מקרא מועט והלכות מרובות", חגיגה א, ט).

דוגמה להתפלגות הכתיב:

תיבות "כָּאוֹר", "כִּיאוֹר" נכתבו באל"ף רק במשנה, לפי כ"י קויפמן (7 פעמים), ואילו בשאר המקורות (תוספתא, ספרי, מכילתא ועוד) נכתבו בעי"ן: כעור, כעורה, כיעור (21 פעם); גם בספר בן-סירא נכתבו בעי"ן (מכוער).

הפועל "גמא", המצוי במקרא, נכתב כך ונוטה כנחי ל"א-ל"י בכל המקורות התנאיים: משנה, ספרי, מכילתא ועוד (7 עדויות). ואילו בתוספתא נכתב בעי"ן: "גמע" ונוטה כשלמים (15 פעם).

עיקר העבודה היום במפעל "המילון ההיסטורי" מתרכזת באיסוף החומר, החסנתו במחשב, ניתוח דקדוקי של המלים ועיבודן לכדי קונקורדנציות. גם כשיוחל בעבודת חיבור המילון ימשיך המפעל לצבור חומר ולהתעדכן, וישמש מקור קבוע ללימוד ולמחקר.