

הישיבה המאתיים חמישים ושש (רנו)

ביום כ"א בכסלו התשס"א (18 בדצמבר 2000) בשעה 15:00, בבית האקדמיה בירושלים.

הנוכחים: החברים ארי אבנר, משה אזר, שרגא אירמאי, יהושע בלאו, יעקב בן-טולילה, משה בר-אשר (יו"ר), גדעון גולדנברג, אפרים חזן, אברהם טל, אבנר טריינין, יוסף יהלום, ישראל ייבין, קציעה כ"ץ, גד בן-עמי צרפתי, מנחם צבי קדרי, נעמי שמר;
החברים-היועצים דוד טלשיר, אהרן ממן, משה עסיס, אלישע קימרון, מנחם קיסטר;
המנהל הכללי נתן אפרתי; עובדי המזכירות המדעית אורלי אלבק, רונית גדיש, עינת גונן, ברק דן, צביה זמירי, רחל סליג; סמדר ברק ושרית גורמס ממכון מזי"א;
מרצה אורח: פרופ' שלמה יזרעאל מאוניברסיטת תל-אביב.

סדר היום: א. הרצאת שלמה יזרעאל "מאגר העברית המדוברת בישראל"
ב. מילים בשימוש כללי
ג. הוועדה למינוח טכנולוגי בחיפה

א. הרצאת שלמה יזרעאל "מאגר העברית המדוברת בישראל"
(הרצאה שנחית לזכרו של משה ברייטברד)

נכבדי, בראשית דבריי אני מבקש להודות מקרב לב לנשיא האקדמיה, שהזמינני להרצות לפני מליאת המוסד הנכבד הזה על התכנית לכינון מאגר העברית המדוברת בישראל.

שנתיים ימים חלפו מאז החלה העבודה על תכנון המאגר. היו אלה שנתיים של עבודה כבירה, של לימוד ושל התייעצויות; שנתיים שבהן עמל צוות רבגוני על תכנון ועל ניסיונות ראשונים "לגרד את השטח" – כפי שנוהגים דוברי העברית המדוברת בישראל לומר. עתה הסתיים שלב התכנון, ואנו עומדים להתחיל באיסוף החומר למחקר המקדים.

תורשה לי הערה אישית. עיסוקי האקדמיים הובילונו אל מחוזותיו של העולם העתיק, אל כנען שלפני היכון ממלכת ישראל הגדולה ואל הרקע שמתוכו – או אל

תוכו – צמחה העברית. עם זאת בכל אותן שנים ארוכות של עיסוק בכתבים עתיקים לא פסקתי מהתבונן בהתפתחותה של העברית בת זמננו ולקונן על החסר במחקרה. מתוך רקע משולב זה נבעה בי ההכרה העמוקה בצורך בכינון מאגר מידע שיצור בתוכו טקסטים למחקר העברית שלנו. שהרי מחקר הלשונות השמיות, ובראש ובראשונה מחקר העברית, הושגת מאז ומתמיד על מחקרים של טקסטים, וכבר ראשוני החוקרים של הלשונות השמיות, מודקקי העברית והעברית בימי הביניים, התבססו על אוסף מוגדר של טקסטים, על מאגר לשון, על קורפוס.

הגישה הבלשנית המבוססת על מאגרי לשון זכתה לתשומת לב רבה במרוצת עשרים השנה האחרונות. בזכות גישה מחקרית זו תיאור הלשון והסקת המסקנות התאורטיות מתבססים על אמצעים סטטיסטיים ועל בדיקת שימושי הלשון הבאים לידי ביטוי במציאות. ענף בלשני זה זכה שייקרא בשם Corpus Linguistics, ועליי להודות שאינני יודע כיצד ייקרא ענף זה בעברית.

המשימה הראשונה שייטול על עצמו הבלשן היא כמובן כינון מאגר לשון. מאגר לשון הוא אוסף של טקסטים, כל אוסף של טקסטים. ההתפתחות האחרונות במדעי המחשב, השינוי המכריע באפשרויות האחסון וגידול הזיכרון במחשב – שיפרו באופן ניכר את תנאי חקר מאגרי הלשון בעולם כולו. כיום נתפס מאגר הלשון כאוסף של טקסטים ממוחשבים. מאגר לשון טוב הוא מאגר המכיל טקסטים ממוחשבים שנוגמו כדי לתת ייצוג מרבי של שפה או של ניב.

העברית שבה לחיים מלאים במאה העשרים. בשלהי המאה התשע עשרה, עם עלייתה של התנועה הציונית, ובעיקר במאה העשרים, הגיעו לארץ ישראל גלי עלייה גדולים, והם שהחזירו לעברית את חיוניותה כשפה ילידית וכשפת היום-יום. כבר חלפו מאה שנות שימוש בעברית, ועולם המחקר החמיץ הזדמנות פז לתעד את נביעתה ואת צמיחתה של מערכת תקשורת לשונית ענפה וכוללת. העברית נתונה עדיין בתהליך של שינויים מואצים בעקבות גלי העלייה הגדולים והשינויים המהירים בחברה הישראלית ובתרבותה. הלכך כינון המאגר של העברית הישראלית הוא צורך דוחק. צורך זה הוא שהביאנו להתמודד עם המשימה של כינון מאגר מקיף של העברית הישראלית.

קריאות לאיסוף מאגר מדעי של העברית החלו להישמע זה זמן. פרופ' יעקב בן-טולילה תיאר את מאגר הצרפתית ממונטרֶאול וקרא לכונן מאגר של העברית המדוברת. ברשימת הביקורת שלו על ספר הדקדוק של גלינרט *The Grammar of Modern Hebrew*, דקדוק המבוסס על עדותם של שישה אינפורמנטים, קרא פרופ' בלאו להשתית דקדוק מקיף על מאגר נרחב של חלופות כתובות ומדוברות של השפה, ולא על קביעות המבוססות על כשירות לשונית מפי דוברים ילידיים ספורים. לפני יותר מעשרים שנה הגיש פרופ' קדרי הצעה לוועדת האקדמיה לכונן מאגר של העברית בת זמננו, בהציעו להתמקד בסקר לשון הספרות. קדרי חזר על קריאה זו בהצאתו לכבוד שנת הלשון העברית, בהדגישו את הצורך הדחוף לכינון מאגר העברית הספרותית החיה. בנימוקיו כבדי המשקל הדגיש בעיקר את הצורך ביצירת תשתית למחקר. נקודת זינוק טובה לעבר המטרה הזאת היא עבודתו של פרופ' שויקה, השוקד כעת על איסוף מאגר ממוחשב של העברית הכתובה. מאגר בר-אילן לעברית החדשה כלל בסוף שנת 1999 עשרים ושישה מיליון מילים.

היות שמאגר שיטתי של העברית הספרותית נמצא בשלבי איסוף, נראה לנו שמוטב להשלים את הנחוץ דווקא בתיעוד החלופות של העברית המדוברת. בניגוד לטקסטים מן העברית הספרותית טקסטים מדוברים הם בני חלוף, ואינם בבחינת קיים הניתן לאיסוף בכל עת. מסיבות אלה נתנו עדיפות ראשונה לתיעוד העברית המדוברת.

מאגר מידע הוא תנאי כל יעבור לתכניות אחרות שבלעדיו אינן בנות השגה; יהא זה דקדוק של העברית החדשה, מילון מקיף או כל תכנית מחקר או פיתוח אחרת לצורך עיוני או מעשי. אפשרויות המחקר הגלומות במאגר שיטתי של שפה הן עצומות וכוללות יישומים בתחומי הבלשנות, החינוך, התרבות והחברה, וכן יישומים טכנולוגיים רבי עצמה.

מטרות המפעל

1. יצירת מאגר של העברית הישראלית המדוברת כתשתית למחקר שיטתי. מחקר המאגר יקיף מגוון רחב של נושאים הקשורים בשפה העברית ובמתודולוגיה הכללית של חקר הלשון המתבסס על מאגרי לשון.

2. הפצת המאגר לציבור במולטימדיה ובדפוס. ההפצה באמצעים אלקטרוניים, כגון תקליטורים ואינטרנט, תיעשה כך שההקלטות ותמליליהן יוצגו במקביל ובשילוב דרכי ניתוח ותיעוד נוספות. אני מבקש להדגיש כי המאגר יהיה נגיש ופתוח לכול, ועם כינונו יופץ חינם או במחיר העלות לכל דורש.

מהות המאגר

המאגר ייצג את מגוון חלופות הלשון של העברית המדוברת בישראל היום. בכונתנו לכלול מדגם מייצג של חלופות דמוגרפיות ושל חלופות תלויות הקשר. החלופות הדמוגרפיות מזהות עם קבוצות דוברים שונות – גאוגרפיות, אתניות וחברתיות; חלופות תלויות הקשר הן פונקציה של מצבים שונים, כגון נושא השיח, בני השיח וכיו"ב.

בכינון המאגר הבאנו בחשבון את המבנה המיוחד של החברה הישראלית. בקרב האוכלוסייה הדוברת עברית מספר הדוברים הילידיים והלא-ילידיים שווה. עובדה זו – עם המורכבות שיצרה ההיסטוריה המיוחדת של העברית החדשה – היא בעלת חשיבות מכרעת להרכב הטקסטים שייכללו במאגר. החלטנו לכלול טקסטים של דוברים ילידיים ושל דוברים לא-ילידיים כאחד. לישראל אישים בעלי השפעה שאינם נמנים עם הדוברים הילידיים, כגון הסופר סמי מיכאל, ראשי הממשלה לשעבר שמעון פרס ויצחק שמיר ורבים אחרים. אוכלוסיית ישראל מוסיפה להתרחב הודות לזרימה המתמדת של העולים אל ארץ ישראל. אף זה גורם בעל השפעה על תגודות במערכת הלשונית, וחשוב לתעדו. תיעוד לשונם של הדוברים הילידיים בלבד לא ישקף את העברית בת זמננו ואת המצב החברתי-הלשוני בישראל, על כל מורכבותו. אילו התעלמנו מהדוברים הלא-ילידיים, היה המאגר מעוות מבחינת ייצוג המערכת הלשונית בישראל, והייתה נפגמת האפשרות לבסס מחקר לשוני או חברתי-לשוני כלשהו על נתונים בני משמעות.

המאגר ישקף מצב סינכרוני של העברית המדוברת בישראל. היות שמדובר בשפה טבעית, ובמיוחד בעברית הישראלית המשתנה בתכיפות, יש להשלימו בפרק זמן קצר ככל האפשר.

מאגרי הלשון הממוחשבים העומדים לרשות החוקרים בעולם שונים בגודלם. במאגר הכולל טקסטים כתובים ומדוברים כאחד חלקם של אלה הכתובים עולה בדרך כלל כדי שני שלישים מכמות הטקסטים הכללית במאגר, וודאי שיש בכך עיוות התפוצה הסטטיסטית של שני הסוגים האלה במציאות. אפילו ב"עם הספר" הלשון המדוברת נפוצה בהרבה מזו הכתובה. עיוות זה מובן מאליו, ואולי אף בלתי נמנע, משום הקלות היחסית שבאיסוף טקסטים כתובים. מאגר העברית המדוברת בישראל עתיד לכלול טקסטים מדוברים בלבד, ושאיתנו היא ליצור מאגר גדול יחסית, כדי לייצג נאמנה את המצב הלשוני העכשווי של העברית בישראל.

אפשר להציג מאגר לשוני בצורות שונות: בהקלטות מתוכנתות, בתעתיק פונטי, בצירוף גלוסות ובתרגום. אפשר להוסיף לו תיוגים שונים: לפי מיון תחבירי, מורפרי סינטקטי, מורפולוגי, מורפרי-פונולוגי. אפשר להקליטו בקלטות שמע ובקלטות וידאו. לכל דרך רישום יתרונות וחסרונות משלה. לדוגמה: קלטת הוידאו יכולה להציג גם מאפיינים חוץ-לשוניים, אך יש בה כדי להגביל אינפורמנטים בטבעיות דיבורים. המאגר יכול את הדברים האלה: הקלטות דיגיטליות בקלטות שמע, תמלילי כל הטקסטים שבמאגר בכתיב עברי, תעתיק פונטי של קטעים נבחרים, ניתוח מורפולוגי שלהם ותרגומם לאנגלית.

ממדי המאגר

מאגר העברית המדוברת בישראל יכול כחמישה מיליון מילים, ובו אלף תאים בני חמשת אלפים מילים כל אחד. מספר התאים וכן מספר המילים שייכללו בכל אחד מהתאים ייקבעו בהתאמה לייצוגם האמתי באוכלוסייה. מאגר של חמישה מיליון מילים בשפה המדוברת נראה גדול דיו לשקף הן את המבנה הכללי הן את המאפיינים הייחודיים של רוב החלופות הלשוניות. רבים מהמאגרים הקיימים כוללים הרבה פחות מחמישה מיליון מילים ואינם מאגר לשוני מייצג. מאגרים גדולים יותר התייחסו לשאלת הייצוג רק באופן חלקי. כינון מאגר של שפה מדוברת שיכלול יותר מחמישה מיליון מילים אינו מעשי. קל לאסוף טקסטים כתובים תוך שימוש בטכניקות סריקה או בחומרי אינטרנט. אך טקסטים מדוברים קשים הרבה יותר לאיתור, להגדרה ולהפיכתם נגישים לשימוש, וזאת בשל המורכבות הרבה שבליקוטם ובתמלולם.¹

המדגם המייצג

על מאגר לשון מייצג להציג שני סוגי חלופות לשוניות: חלופות דמוגרפיות וחלופות תלויות הקשר. לפיכך עלינו לאסוף נתונים על פי שני סוגי קריטריונים אלו. מאגר העברית המדוברת בישראל יורכב משני תת-מאגרים: מאגר ראשי ומאגר משלים. המאגר הראשי יהיה חלק הארי של המאגר כולו ויכלול כתשעים אחוז מכלל הנתונים; המאגר המשלים יכלול כעשרה אחוזים מהנתונים. המאגר הראשי יורכב מתאים שייבנו על פי קריטריונים דמוגרפיים משולבים בקריטריונים לבחינת הקשרי השיח. הנתונים במאגר המשלים ייערכו בשני תת-מאגרים שונים: האחד יועד להשלמת

1. [הצרת העורך: במונחי האקדמיה המונח הוא תְּכַתּוּב.]

המאגר הראשי ויושנת בעיקרו על קריטריונים דמוגרפיים. תת-מאגר זה ידגום בדגימה לא מאוזנת קבוצות אוכלוסין שייצוגן יהיה חסר במאגר הראשי; תת-המאגר השני יושנת בעיקרו על קריטריונים של הקשרי שיח ויכול דגימות טקסטים מאמצעי התקשורת, מן הכנסת ומבתי המשפט. כל תת-מאגר מן המאגר המשלים יורכב מחמישים תאים, דהיינו חמישה אחוזים מן המאגר כולו.

המאגר על שני חלקיו, הראשי והמשלים, יאוחסן בבסיס נתונים רחב ומתוחכם. כל חוקר יוכל לשלוף את הנתונים כחפצו לפי שילובים מגוונים של משתנים. לדוגמה, הוא יוכל להתמקד במאפייני לשונם של דוברים ילידיים ממוצא צפון-אפריקאי, צעירים ובעלי השכלה על-תיכונית.

בהרצאתי זו אני מעוניין להציג את התכנית לכינון המאגר הראשי, הוא המאגר המייצג. אולם לפני כן ראוי להעיר כמה הערות נוספות על דרך השימוש במאגר.

חשובה ההבחנה בין הקריטריונים שישמשו לדגימה ובין הקריטריונים שישמשו לניתוח החומר הלשוני. האתגר העיקרי בתיעוד מדגמי של אוכלוסייה בת מיליונים לצורך מחקר שימושי הלשון וחלופותיה הוא להגיע לייצוג רחב ככל האפשר של חלופות הלשון, לתעד טקסטים באורך מספיק לחקירתם ועם זאת לשמור על ממדים סבירים לעיבוד הנתונים במאגר. המשימה מורכבת עוד יותר בשל העובדה שהנתונים הדמוגרפיים כשלעצמם נערכים על פי רובד נוסף, רובד הקשרי השיח על מאפייניו השונים. בעוד שהמאגר יכיל בבסיסו נתונים טקסטואליים שייאספו על פי בדיקה מדגמית אקראית של שכבות האוכלוסייה, מידע דמוגרפי מפורט ומידע על הקשרי השיח יילקט עם ביצוע ההקלטות, יאוחסן במאגר המידע ויאפשר עיבוד סוציו-לינגוויסטי ולשוני מורכב בהרבה.

דגימת האוכלוסין עצמה תיעשה על פי קריטריונים סטטיסטיים כדי לשקף כמותית את כלל האוכלוסייה. לעומת זאת איסוף נתונים על פי קריטריונים אנליטיים לא בהכרח יהיה מייצג מבחינה סטטיסטית. במילים אחרות, המשתמש בבסיס הנתונים של המאגר יוכל לשלוף נתונים לשוניים מבסיס נתונים שאמנם יהיה ערוך על פי קריטריונים סוציו-לינגוויסטיים, אולם לא בכל מקרה יהיה בנתונים משום ייצוג כמותי של האוכלוסייה שהטקסטים האלה מייצגים.

לפיכך נראה לנו לנקוט קריטריונים שונים בעת הדגימה מאלה שיינקטו בעת ניתוח הממצאים. בעת עריכת המדגם יילקחו בחשבון הנתונים הדמוגרפיים, ואלה יידגמו על פי שיטות הכימות המקובלות. עריכת מדגמים היא תחומם של סטטיסטיקאים ושל סוציולוגים. תהליך עריכת המדגם – עם כל חשיבותו בעת תכנון המאגר – אינו רלוונטי למשתמש. לעומת זאת הקריטריונים שיינקטו בעת העיבוד המדעי של הטקסטים, דהיינו על ידי החוקרים שיעשו שימוש בנתונים, חייבים להיות כפופים לקריטריונים הנהוגים במחקרים בלשוניים וסוציו-לינגוויסטיים. הקריטריונים האנליטיים לגבי הנתונים הסוציו-לינגוויסטיים במאגר יהיו שונים ורבים יותר מאלה שישמשו אותנו בעת כינון המאגר וארגונו. כדי להקל על המשתמש ולאפשר הסקת מסקנות הנשענות על בסיס איתן, יהא עלינו להגדיר כמה קריטריונים אנליטיים, הן דמוגרפיים הן בלשוניים, עוד בשלב ארגון החומר כדי לשמר את עקרון הייצוגיות.

המאגר הראשי

לשם כינון המאגר המייצג נקטנו גישה תלוית תרבות. תבנית המאגר הראשי תוכננה כך שתתאים למבנה הייחודי של החברה הדוברת עברית. המאגר שואף לגשר בין איך-ספור חלופות השיח הנהוגות בקרב קהילת הדוברים של העברית הישראלית ובין ייצוגיותה במאגר על ידי אפיון השונות במונחים דמוגרפיים והקשריים גם יחד. למיטב ידיעתנו, זהו ניסיון ראשון וייחודי בעולם כולו לכונן מאגר מייצג על שני צירי משתנים – דמוגרפיים ותלויי הקשר – בהתאם לקריטריונים סטטיסטיים ואנליטיים. כל אחד משני הצירים הוא מרחב רב-ממדי בעצמו. המודל האנליטי של המאגר יכול ארבעים וחמישה תאים דמוגרפיים אשר כל אחד מהם יכול שמונה תאים שונים על פי דגימת ההקשר. כל תא מתאים אלו עשוי להכיל בעצמו ארבעה תאים.

“תא” הוא היחידה הסוציולוגוויסטית הבסיסית במאגר. התא הוא קטע דיבור מוקלט בן חמשת אלפים מילים של טקסט רציף ומלוכד או כמה קטעים שכוללים יחדיו חמשת אלפים מילים. המאגר הראשי יכול תשע מאות תאים שיידגמו בדגימה אקראית. במסד הנתונים ייאגר מידע סוציולוגוויסטי על אודות דוברים שונים: מקום היוולדם, היותם דוברים ילידיים או לא-ילידיים, עדה, מקום מגורים, סוג היישוב, גיל, מין, מעמד חברתי-כלכלי, מקצוע, עיסוק, שירות צבאי, דת, שהייה ממושכת בחוץ לארץ ושפות המדוברות בבית. דגימה על פי הקריטריונים האלה כולם תנפח את המאגר לממדים עצומים, וכאמור כינון מאגר גדול מחמישה מיליון מילים אינו מעשי. מסיבה זו ניצמד בעת דגימת האוכלוסייה להיקפים סטטיסטיים מקובלים. עם כל אחד מן האינפורמנטים שיוקלטו ייערך ריאיון כדי להפיק את מרב המידע הסוציולוגי הרלוונטי עליו, מידע שישמש מחקרים בלשניים וסוציולוגוויסטיים נוספים. כל הקריטריונים שפירטנו לעיל יכללו בבסיס הנתונים הסוציולוגוויסטי ויהיו נגישים לכל דורש. מחלוקת יתר של המדגם לתת-קבוצות רבות מדי יתקבל אוסף של אידיולקטים, ולא דווקא מאגר מייצג ונגיש של כל קהילת הדוברים. על כן נצמצם את מספר המשתנים שלפיהם יאורגנו נתוני המאגר, כדי לאפשר אחזור מידע מייצג בר-כימות של דוברי העברית.

הנחת העבודה שלנו מתבססת על שלושה קריטריונים דמוגרפיים שנראים לנו החשובים ביותר ביצירת השונות הלשונית בישראל:

- (א) מקום לידה, ארץ מוצא המשפחה, עדה או דת. קריטריון זה נחלק לחמש קטגוריות: (1) יהודים, ילידי הארץ, בני אב יליד אסיה או אפריקה; (2) יהודים, ילידי הארץ, אחרים; (3) יהודים, ילידי חו"ל שעלו לפני 1965; (4) יהודים, ילידי חו"ל שעלו אחרי 1965; (5) לא-יהודים (מוסלמים, נוצרים, דרוזים);
- (ב) גיל. קריטריון זה נחלק לשלוש קטגוריות: (1) צעירים (בני 15–27); (2) בוגרים (בני 28–50); (3) מבוגרים (בני 51 ומעלה);
- (ג) השכלה. קריטריון זה נחלק לשלוש קטגוריות: (1) מי שלא סיים בית ספר תיכון; (2) בעלי השכלה תיכונית; (3) בעלי השכלה על-תיכונית.

אם כן הציר הדמוגרפי של תבנית התאים הרב-ממדית הוא עצמו רב-ממדי וכולל שילוב הקטגוריות של משתני מקום הלידה, המוצא והעדה או הדת עם משתני הגיל ועם משתני ההשכלה. קריטריון משמעותי נוסף הוא מין הדוברים. תבנית התאים המוצעת

אינה כוללת נתון זה כקריטריון נפרד משום שמין הדוברים יובחן בכירור בעת הדגימה, ואנו צופים כי היחס בין גברים ונשים בכל קטגוריה אנליטית יהיה שווה. כדוגמת הציר הדמוגרפי, גם ציר הקשרי השיח מעוצב כתבנית רב-ממדית. בעת קביעת החלופות תלויות ההקשר הבאנו בחשבון חמישה משתנים, שלושה עיקריים ושניים משניים. במשתנה (א) משתקפים היחסים האישיים: יחסים בין בני משפחה או בין חברים קרובים יוגדרו כיחסי קרבה; במשתנה (ב) משתקף מבנה השיח: מבנה שיח שבו חלוקת תפקידים ברורה בין בני השיח (למשל אם למי מהם תפקיד סמכותי) יובחן משיח שאין בו חלוקת תפקידים כזאת; משתנה (ג) יבחין בין נושא שיחה אישי ללא-אישי. במשתנים המשניים משתנה (ד) יבחין בין מונולוג לדיאלוג. אף שמונולוג כשיח של אדם יחיד עשוי להימצא בכל אחת מההקלטות, נבחין במשתנה זה רק במקרים המובהקים שבהם מתפתח מונולוג, כלומר כשהמונולוג מגדיר מהותית את סוג השיח; המשתנה האחרון (ה) יבחין בין שיחה בטלפון לשיחת פנים אל פנים. שלושת המשתנים העיקריים – (א), (ב), (ג) – יבואו לידי ביטוי במאגר בכל שמונת הצירופים האפשריים שלהם. לעומת זאת המשתנים המשניים – (ד), (ה) – ייושמו רק בחלק מהתבנית בהיותם הרבה פחות שכיחים ברוב קהילות הדוברים. אנו מניחים כי הן מן הבחינה התאורטית והן מן הבחינה המעשית אלה הם המשתנים המבחינים בין סוגי השיח העיקריים. יחידות שיח שיאובחנו על פי משתנים אלו ישקפו מבנה לשוני שונה. את ההנחה הזאת נבדוק במחקר המקדים ובמהלך איסוף הנתונים.

הדגימה

מטרת כינון תאים מייצגים מבחינה סטטיסטית היא לאסוף נתונים מספיקים למחקר סוציו-לינגוויסטי או בלשני. נשתדל לכוון מאגר שגם ישקף את המערכת הלשונית בכללה וגם ייצג נאמנה מגוון רחב של שונות לשונית. עריכת מדגם מייצג על פי נתונים דמוגרפיים היא טכניקה שכיחה הנהוגה בדגימת אוכלוסין ונעשית על ידי דגימה אקראית של האוכלוסייה, למשל על פי מקום מגורים. לעומת זאת הניסיון שנצבר בבניית מאגר לשון שייצג נאמנה גם את הקשרי השיח מועט ביותר. כדי להבטיח ייצוג אמין של השונות הלשונית, הדמוגרפית ותלויות ההקשר כאחת, יישמר עקרון האקראיות בעת דגימת כל הנתונים הטקסטואליים. דגימת הקשרי השיח תיעשה לאחר איסוף כל ההקלטות של האוכלוסייה הנדגמת דגימה דמוגרפית. כל מי שעלה במדגם האוכלוסין האקראי יתבקש להקליט את כל מהלכיו והקורות אותו במשך יממה רצופה אחת. כל הקלטה כזאת תיבדק ויוסרו מרווחי שתיקה וקטעי דיבור לא ברורים. מן החומר הנקי ייבחר באקראי קטע באורך של שעה, והוא יהיה מקור הנתונים שייכלל במאגר הראשי. בחירה אקראית זו עשויה לשקף את ביזור מצבי השיח הטבעיים הן על פי היקריותיהם בזמנים שונים, הן בהתאם להקשרי השיח והן בהתאם לסוגי השיח המזומנים לסוגי אוכלוסייה שונים.

גישור הפער

בשל אקראיותה והיקפיה של דגימה זו היא אינה תואמת בהכרח את המודל התאורטי שעל פיו נבנתה תבנית התאים. נצטרך אפוא לתת את הדעת על גישור הפער בין תביעות הסטטיסטיקה לקריטריונים האנליטיים. התכנית הבסיסית של המאגר מקנה עשרים תאים

לכל חלופה דמוגרפית. היות שיש ארבעים וחמש חלופות דמוגרפיות, יצטברו משילובן תשע מאות תאים. לא כל החלופות תואמות זו את זו. למשל, קריטריון מקום הלידה השלישי (יהודי שאינו יליד הארץ ועלה לפני 1965) אינו תואם את קבוצת הגיל הראשונה (15–27 שנים) ואף לא את כל מרחב קבוצת הגיל השנייה (28–50 שנים), שגם בה ילידי התקופה שלאחר 1965. נוסף על כך, חלק מהנדגמים בקבוצת הגילים הצעירים אינם יכולים להיות בעלי השכלה על-תיכונית. הלכך, כבר על פי התכנון הראשוני לא תוכל התבנית להתמלא בכל תשע מאות תאיה. כמו כן אנו סבורים שלא כל החלופות האפשריות מתוך החלופות תלויות ההקשר זמינות וקיימות עבור כל אחת מן החלופות הדמוגרפיות.

כבר הבהרתי כי איסוף הנתונים לשם הצבתם במאגר וליקוטם בעת השימוש המחקרי בהם יתבצעו בדרכים שונות לחלוטין. ראשית יתנהל איסוף הנתונים להצבתם במאגר. משימה זו תבוצע באמצעים סטטיסטיים מוקפדים, כלומר על פי דגימה אקראית. ארגון החומר הטקסטואלי שנאסף הוא השלב הבא. עריכת הנתונים לשימוש מחקרי תבסס על העקרונות שנקבעו בכינון תבנית התאים, ותתייחס לקריטריונים הדמוגרפיים ולקריטריונים תלויי ההקשר ששוקללו, כפי שתיארתי לעיל. לרשותנו יעמדו טקסטים שנבחרו באקראי מתוך מצאי טקסטים שהפיקו אנשים שנבחרו באקראי במגוון רחב של הקשרי שיח. הטקסטים הללו ייערכו בתבנית הרב-ממדית, וכל יחידת טקסט תוכנס אל התא התואם לה על פי הנתונים הדמוגרפיים של דובריה ועל פי הקשר השיח המתאים. חלוקת החומר לתאים תיעשה על פי עקרון האקראיות ובכימות הולם. דהיינו, תבנית התאים תתמלא במצאי הטקסטים שהבחירה האקראית תעמיד לרשותנו ולא על פי מספר התאים שהעיצוב העקרוני יעמיד לרשות כל יחידה דמוגרפית. גישה זו תאפשר לנו לעמוד על מגוון דגמי השיח הקיים בקהילת הדוברים וללמוד על הדגמים מבחינה כמותית ואיכותית, כפי שיעלה משילוב הקריטריונים הדמוגרפיים והקריטריונים תלויי ההקשר. במילים אחרות, המיון לתאים יניב ניתוח ראשוני של הקשרי השיח שבשימוש חלקי אוכלוסייה שונים, וכן ניתוח ראשוני של היקף השימוש של החלופות הלשוניות הללו ביחס לחלופות אחרות. כפי שציינתי, כל הנתונים הסוציולוגיים וסטיים יהיו נגישים לכל דורש, וחוקר המאגר יוכל לשלוף טקסטים לצרכיו על פי נתונים סוציולוגיים כחפצו. עם זאת כדי להניב תוצאות משמעותיות בניתוח לשוני, אנו מציעים משתנים כוללים יותר כבסיס לשליפת נתונים.

פתחתי בהערה אישית ואסיים בהערה אישית. המפעל לכינון מאגר העברית המדוברת בישראל הוא ללא ספק מפעל כביר, והבלשן שבי אף יאמר מלהיב. אולם לעניין כינון המאגר אינני פועל אך ורק כבלשן, אלא אף ובעיקר כדובר עברית, כישראלי וכיהודי, כבן עם היונק מפאָר עברו ובונה את עתידו. בניית העתיד – בחינוך יסודה וקימומה. למאגר העברית המדוברת בישראל יישומים בתחומים רבים. לגבי דידי, יישומי החינוכיים עולים על כולם. הכרת הלשון המדוברת ותפוצת חלופותיה תהא בסיס ראשון במעלה להכרת החברה בישראל ולהבנת תרבותיה. התבוננות בשונות הלשונית תישא עמה הכרה עמוקה בצורכי החברה על מגוון רבדיה, וגם לימוד הלשון לא ייזק. נדמה לי כי הניסיון מוכיח שהחינוך הלשוני בארץ אינו נושא פרי הילולים. בלי הכרת הנתונים לא

ייתכן מחקר, ובלי מחקר לא ייתכן שינוי. חקירת הלשון המדוברת תתרום לא רק לשינוי בדגמי החינוך הלשוני הקיימים היום, אלא תביא למהפכה של ממש בהבנת הצורך והמהות של החינוך הלשוני ושל ההשלכות מרחיקות הלכת שמהפכה כזאת נושאת בחובה. והלוואי שיינתן לידינו לקרב ולו אך במעט את מאמר הנביא ישעיהו: "וְלִבְבֵי נְמֻקִּים יִבִּין לְדַעַת וְלִשׁוֹן עֲלֵגִים תִּמְהַר לְדַבֵּר צְחוֹת" (לב, ד).

לא אצא ידי חובתי אם לא אודה לשותפיי למחקר, לחברי הצוות שנלהבו כמוני לצאת לדרך ארוכה זו של כינון המאגר: בנימין הרי מאוניברסיטת אַמְוֹרִי באטלנטה ומאוניברסיטת תל-אביב; ג'ון דו-בואה מהאוניברסיטה של קליפורניה בסנטה ברברה; גיורא רהב, מירה אריאל, אליעזר בן-רפאל ואילנה שוהמי מאוניברסיטת תל-אביב; חבר האקדמיה יעקב בן-טולילה מאוניברסיטת בן-גוריון בנגב; אוטו יסטרו מאוניברסיטת ארלנגן-נירנברג בגרמניה; שמואל בולוצקי מהאוניברסיטה של מסצ'וסטס באמהרסט; וג'פרי כאן מאוניברסיטת קמברידג' באנגלייה.

פרטי התכנית לכינון מאגר העברית המדוברת בישראל הועלו באינטרנט במטרה לפרסם את תכניתינו ברבים וכדי לבקש תגובות ועצה טובה מן המתעניינים בהן. נודה על כל תגובה והערה.

[הערת העורך: ראה גם ש' יורעאל, ב' הרי וג' רהב, "לקראת כינון מאגר העברית המדוברת בישראל", לשוננו סד (התשס"ב), עמ' 265-287.]

מ' בראשור: תודה רבה לאד' יורעאל. נראה כי מקובל על כולנו שאי אפשר להפריז בחשיבות המאגר הקורם עור וגידים. אם יש שאלות או הערות על המאגר, הרשות נתונה. נרכז תחילה את השאלות ואת התגובות.

מ' עסיס: המשימה היא גדולה. אינני יכול להעריך כמה זמן וכמה מאמץ ידרשו להקלטת חמשת אלפי התאים שיאצרו חמישה מיליוני מילים. אבל כשאני משווה מאגר זה למאגר הספרותי של אוניברסיטת ברא-אילן, הכולל כבר היום עשרים ושישה מיליון מילים, דומני שכדאי לתכנן שלב שני של חמישה מיליוני מילים נוספות.

ד' טלשיר: מתי לדעת האד' יורעאל אפשר יהיה להתחיל ליהנות מפירות המאגר?

מ' צ' קדרי: הצגת הדברים הייתה מרשימה, מפורטת, ואף אפשר לומר מלהיבה. האם חשבתם על הקמת גוף מייעץ, ציבורי ומדעי, שילווה את המפעל? אני יודע שבהכנת החומר המילוני של מפעל קטן בהרבה, המילון רב-מילים, הייתה התייעצות מרובה בכמה וכמה משלבינו. אחת השאלות המרכזיות הייתה מה ייכנס למילון ומה לא.

א' אבנר: לא מכבר התנהל באקדמיה ויכוח בדבר תפוצת הצורות פֶּעֶלֶן-פֶּעֶלֶן במילים מסוימות. האם יוכל המאגר לתת מענה על שאלות כאלה?

א' טל: תהיתי איך נקבעו הקריטריונים לעברית מדוברת. יקרה למשל שאדם יעלה משלב כדי לצטט במהלך דיבורו או סיפורו.

א' ממך: ברשימת המזכרים בסוף דברי האד' יזרעאל נזכרו אנשים מכל רחבי תבל. מדוע נפקד מקומו של נציג מן האוניברסיטה העברית?

ש' יזרעאל: אפתח בשאלה האחרונה. תחילתו של המפעל ברעיון שהתחלנו, אני ובנימין הרי, לגלגל. לאט לאט הצטרפו אנשים על פי תחומי התעניינות וניסיון מקצועי. רק מקרה הוא שאין בשותפים למפעל איש אוניברסיטה זו או אחרת.

אשר לשאלתו של האד' עסיס, יש הבדל גדול מאוד בין בניית מאגר של שפה כתובה למאגר של שפה מדוברת. מאגר בן חמישה מיליוני מילים נחשב גדול מאוד. מן המאגרים הידועים לי רק *British National Corpus*, שהושקעו בו מאמצים ומשאבים יוצאים מגדר הרגיל, גדול הימנו – בן עשרה מיליוני מילים. היקפם של כל יתר מאגרי הלשון המדוברת שאני מכיר אינם עולים על מיליון מילים.

את ניצני פרות המחקר אני מקווה שאפשר יהיה לראות תוך שנה עד שנה וחצי; בתחילה במסגרת מאגר חלוץ. כינון המאגר כולו יארך בוודאי שנים מספר, ואם נתגבר על כל המכשולים הטכניים, הארגוניים והכספיים, מדובר על פרק זמן של לפחות שבע שנים מהיום.

גוף ציבורי מיעף אין למפעל, אבל צוות הליכה – שאנשיו באים מן האקדמיה – מתפקד גם כצוות מבקר. מכל מקום אנחנו שמחים לכל הערה והשגה מכל גוף או אדם שהוא, וכבר הכנסנו לא מעט שינויים שמקורם בהצעות שקיבלנו מן החוץ.

מאגר השפה המדוברת איננו מילון השפה המדוברת. לפיכך כל מה שיוקלט ייכנס למאגר, אלא אם יימלך אדם בדעתו ותוך 48 שעות מן ההקלטה יבקש שתימחק הקלטתו. מכל מקום פרטי הדוברים יישארו חסויים.

על שאלתו של האד' אבנר אני יכול להשיב בחיוב. אפשר יהיה להוציא מן המאגר נתונים על תפוצתן של צורות. כמובן תידרש תכנת תיוג מתאימה לניתוח מורפולוגי מדוקדק של כל מילה ומילה.

שאל האד' טל איך תתאים עברית שאינה מדוברת למאגר של עברית מדוברת, ואני מודה לו על השאלה. הזכרתי את המשתנים הדמוגרפיים ואת משתני ההקשר (קרבה, תפקיד, אישי, מונולוג, טלפון). על פי משתני הקשר השיח אפשר לאפיין את השיחות השונות ולהכניסן לתא הנכון. גם להרצאה שנישאת מן הכתב אנו מתייחסים כאל חומר מדובר, אך היא תיכנס לתא שונה מזה שאליו תוכנס שיחת חולין של בני משפחה במטבח, למשל. שימו לב שהקשרי השיח אינם משלבים. איננו מתייחסים לאוצר המילים, לתחביר, לצורות, לאינטונציה וכדומה. תפיסת המשלבים השתנתה בשנים האחרונות בעקבות מחקרים, בעיקר על מאגרי לשון באנגלית, והיום מדברים על טיפוסים טקסטליים. כעת אנחנו בעיצומו של מחקר מקדים לבדיקת ההנחות שלנו על הקשרי השיח. למשל, ייתכן שנמצא שאוצר המילים, התחביר והשימוש בצורות בשיחת טלפון זהים לאלה של שיחת יום-יום במטבח, ואז אין טעם להפריד בין שני טיפוסים אלו.

מ' בראש: תודה רבה לפרופ' יזרעאל. אני מקווה שגם הדברים התאורטיים וגם התוצאות הראשונות – ולו מקצתן – יראו בקרוב אור בכתובים, ויפתח דיון על הצד המעשי של המפעל, על השלכותיו ועל יישומו.

ב. מילים בשימוש כללי

הודעת הוועדה למילים בשימוש כללי

1. מְמַעֵד – המילה תשמש הן לכלי המועד / מרסק / כותש שום, הן לכלי המועד תפוחי אדמה ושאר ירקות. לפי הצורך יתווסף סומך, כגון ממעך שום, ממעך תפוחי אדמה.
2. אֶחָסָן – בּוּיָדָם
3. אֶזוֹר תְּפִעוּל – ground support area
4. רְפָאֵי חָרוּם – פרמדיק
5. (מְקָצוּעַ) נְלִיָּה רְפוּאָה – (מקצוע) פרה-רפואי, allied health professional
6. מְקָלוֹר – stick-light
7. צִחְלוּלִים – زغردة (ברבים: زغاريد; קולות שמחה המקובלים בתרבות המזרח)
8. צְמוּדוּזִים – "טייטס" (מכנסיים צמודים)

רקע

1. מְמַעֵד
במונחי כלכלת הבית (התשל"ח) לא נמצא מונח לציון הכלי (הידני) למעיכת תפוחי אדמה; הכלי למעיכה או ריסוק של שום נקרא מְמַעֵד. במילונים אבר־שושן, החווה, רב־מילים מְמַעֵד הוא הכלי למעיכת תפוחי אדמה. בחנויות ובספרי בישול רווח הצירוף פוֹתֵשׁ שוֹם לציון הכלי למעיכה או לריסוק של שום, ומוֹעֵד (תפוחי אדמה) לציון הכלי למעיכת תפוחי אדמה. הצעת הוועדה – לקבוע מְמַעֵד לשני הכלים. בכך לא נשנה החלטה קודמת של האקדמיה (מְמַעֵד – למעיכת שום) ולא נסתור את המצוי במילונים (מְמַעֵד – למעיכת תפוחי אדמה).

2. אֶחָסָן – בּוּיָדָם
מדובר בייעלית גג או תא מעל תקרת הבית המשמשים בעיקר לאחסון חפצים" (על פי מילון רב־מילים). המונח בגרמנית: Boden; למילה זו שתי הוראות: רצפה; רצפת חלל הגג. המונח בצרפתית: entresol – שתרגומו המילולי 'בי־קומתי'. הצעות אחרות שעלו בוועדה: מחסן עילי, עליית תקרה, עלייה, עליית קיר, תְּקָרִיָּה (מילה משמשת). בנוגע להצעה עֲלִיָּה הוער כי מילה זו מציינת במקורות מקום לגור בו.

3. אֶזוֹר תְּפִעוּל – ground support area
האד' רפי אלבו, מהנדס ראשי למבנים ותשתיות במסגרת פרויקט נתבי"ג 2000, הסביר שמדובר בשני אזורים ממזרח לטרמינל וממערבו – שישמשו לפעילויות התפעוליות והתחזוקתיות התומכות בפעילות הטרמינל, ובהן רחבות לתחזוקת המטוסים, ניטול הכבודה, מחסני ה"דייט־פרי", מפעלי מזון, מפעל לכביסה, מוספים, האנגרים למטוסים, תחנה לתחזוקת אוטובוסים, תחנות דלק תפעוליות, משרדים עורפיים של חברות התעופה, חדרי אוכל לעובדים וכיו"ב.

שתי ההצעות שחעלה הפונה: אֶזוֹרֵי תְּמָד, אֶזוֹרֵי תְּפִעוּל. בוועדה הוצע גם מְתַחְמֵי תַמֵד / תְּפִעוּל; מְרַכְזֵי תַמֵד / תְּפִעוּל. הועלתה גם ההצעה מְתַמָד. הובעה הדעה שאין צורך של ממש בהכרעה של האקדמיה בעניין זה. כנגד זאת נאמר שקביעת האקדמיה תתרום לאחידות במינוח. בשיחתנו עם הפונה הוא העדיף את ההצעה אֶזוֹרֵי תְּפִעוּל, והיא התקבלה בוועדה.

4. רפאי חרום – פרמדיק

המונח נדון בעבר בשלושה סבבים בוועדה. לאחר הסבב הראשון הובא המונח רפאי לפני המליאה, וזו דחתה אותו (שיבה קצח, כסלו התשנ"א; זיכרונות האקדמיה לח, עמ' 36–40). בסבב הזה הופיעו לפני הוועדה נציגי הפרמדיקים, ואלו עיקרי דבריהם: במלחמת וייטנאם הותר לאנשים שאינם רופאים לבצע פעולות רפואיות פולשניות בשדה הקרב. שיטה זו נהוגה בכמה ארצות, וכן בישראל משנת 1979. בצה"ל משתמשים במונח 'חובש החיאה בכיר', אלא שלחובש אסור לבצע פעולות פולשניות. הפעולות שהפרמדיקים עושים: החיאה; טיפול בהפרעות קצב, השתלת קוצב לב חיצוני; טיפול במצוקות נשימה; הזרקת תרופות לווריד; טיפול חירום בילדים. למעשה אין הפרמדיקים מרפאים. בתשובה לשאלה אמרו הנציגים כי אין מונח לציון המקצוע עצמו ומשתמשים ב'פרמדיק' לשם כך. לדבריהם, אפשר לומר שהפרמדיק הוא מומחה לרפואה דחופה.

החלטת הוועדה: רפאי חרום. בצורת הרבים המילה 'רפאי' מזדהה עם רפא, מונח המציין את העוסק בשחזור חרסים. בתוספת הלוואי תיפטר הבעיה. אשר ל'רפאי' – כפי שיש 'מהנדסי' ו'הנדסאי' כך יהיה 'רופאי' ו'רפאי'. רופאים ממשד הבריאות שנשאלו על השימוש בשרש רפאי ל'פרמדיק' השיבו (ברובם) שאין מניעה להשתמש בו.

ד"ר יעקב ברטוב, ראש מגמת "רפואת חירום" (להכשרת פרמדיקים לתואר ראשון) מהפקולטה לרפואה בבאר-שבע, תמך בהצעה רפאי חרום. לדבריו, אפשר למצוא לה תימוכין בשם המגמה 'רפואת חירום'.

5. (מקצוע) נְלוּה רְפוּאָה – (מקצוע) פרה-רפואי, allied health professional

ממשד הבריאות נמסר שמדובר בכעשרים מקצועות פרה-רפואיים, ובהם קלינאות תקשורת, פיזיותרפיה, ריפוי בעיסוק, כירופרטיקה, דיאטנות, אופטומטריה, רוקחות, וכן ברפואת שיניים – שינניות, טכנאי שיניים וסייעות. אפשר לומר שמדובר בכל מקצוע שעניינו רפואה ושהעוסקים בו אינם רופאים. (במשד הבריאות פועלים להסדיר את המקצועות האלה בחוק, ומכל מקום לכל המקצועות האלה יש תעודות הכרה-במעמד הניתנות למי שעומד בדרישות ההשכלה וההכשרה.)

בוועדה עלו ההצעות האלה: מקצוע תְּמָךְ רפואי, מקצוע סְמָךְ רפואי, מקצוע מְצַד-רפואי, מקצוע עזר רפואי, מקצוע נְלוּה רפואה, מקצוע רְפוּאָתִי, מקצוע מְצַר רפואי / מקצוע מְצַרן רפואי. ההצעה מקצוע נלווה רפואה (וברבים: מקצועות נלווי רפואה) התקבלה בוועדה בדיון השני.

6. מְקָלוֹר – stick-light

מדובר בכלי מוארך, דמוי מקל, הפולט אור זוהר, ומשמש בשעות החשכה להכוונת התנועה, לסימון מעברים וכד'. שני פונים בעניין זה הציעו לכוונת את הכלי מקלור. ד"ר דן לינק מן הרשות לבטיחות בדרכים המליץ לאמץ את ההצעה. הוועדה קיבלה את ההצעה פה אחד.

7. צְהַלּוּלִים – زغردة (ברבים: زغاريد)

מדובר בשם לקולות השמחה המושמעים בשמחות – בעיקר מפי נשים מעדות המזרח. את המילה 'צהלולים' הציע האד"י חיים אוליאל משדרות בנימוק שיש בה משום שילוב של 'צהלות', 'יהלולים', וחקול 'לר-לר-לוי'.

יצוין שהמילה 'צהלולי' מובאת במילונים כמילה נרדפת ל'צהלה'. הוועדה החליטה פה אחד לאמץ את ההצעה.

8. צְמוּדוֹנִים – "טייטס" (מכנסיים צמודים)

הפונה האד"י אבי אלון הציע צְמוּדוֹן. הוועדה העדיפה את צורת הרבים 'צמודונים' כדי לקשור את המילה לימכנסיים (שם עצם שאינו ביחיד).

בוועדה עלו גם ההצעות: מְצַמְדִים, הָדוּקִים.

מ' בר-אשר: הוועדה למילים בשימוש כללי הניחה לפנינו שמונה הצעות חדשות פרי עבודתה. נפתח בשלוש הראשונות. האם ליושב ראש הוועדה יש מה להוסיף על הנאמר בחומר הרקע שנשלח לחברי האקדמיה?

מ' קיסטר: על המילה בוידם אני רוצה להעיר כי אף שיש לה יותר משימוש אחד, אנחנו מתבקשים לעסוק רק באחד מהם כפי הגדרתו לפנינו, כלומר רק באותם חללים שהם 'עליית גג או תא מעל תקרת הבית המשמשים בעיקר לאחסון חפצים'.

י' בלאו: באופן עקרוני אני חושב שרצוי למעט בחידושי מילים. על כן אני תומך מצד אחד בהצעה אזור תפעול, שרכיביה אינם חדשים אבל צירופם חדש, ומצד שני מסתייג מן ההצעה אחסן לבוידם. תחתיה יכולה לשמש המילה עלייה, שיתרונה הבולט שאיננה חידוש. אף שלמילה עלייה כמה משמעויות, ההקשר שהדברים נאמרים בו הוא בדרך כלל חד-משמעי, ולכן אינני חושב שהקביעה עלייה תקשה על הדוברים.

א' אבנר: תחילה עליי להעיר שביידיש, שממנה נשאלה לעברית המילה בוידם, אין בוידם אלא 'עליית גג'. לגופו של עניין, מן המילה אחסן גם אני מסתייג. חללי אחסון יש ממינים שונים ובמקומות שונים, למשל מתחת לגרמי מדרגות, ואינני רואה מדוע תיקבע המילה אחסן לחלל אחסון מסוג אחד בלבד.

מ' אזר: גם בעיניי ראוי שהמילה אחסן תציין דבר כללי יותר. אשר לממעך, מאחר שלכלי למערכת שום מקובל השם כותש שום, לא הייתי קובע לו את השם ממעך, אלא את השם המקובל.

א' חזן: אני מציע לבוידם את המילה תְּקָרוֹן, צירוף של תקרה וארון.

נ' שמר: אני רוצה לפתוח בדברי שבח, ועל כן אזכיר תחילה את הצהלולים ואת הצמודונים; הצעות נהדרות. בעניין צהלולים גם התקשרתי לחיים אוליאל והחמאתי לו על הצעתו.

אשר לשלוש המילים הראשונות, ההצעה אזור תפעול יפה בעיניי, ואילו אחסן אינה טובה. לדעתי יש ללכת בכיוון שהציע האד' חזן, כלומר למצוא לבוידם חלופה הבנויה מן המילה תקרה. הצעתו תקרון טובה מאוד.

אולם עיקר הבעיה בממעך. שום כותשים, ולכן אפשר לקרוא לממעך השום פּוֹתְשִׁים. נותרה אפוא בעיית ממעך תפוחי האדמה. פעם הסברתי למי ממשפחתי איך מכינים מחית תפוחי אדמה – הקרויה בפי כול פִּיֶּרָה – ואמרתי שאחרי שמבשלים תפוחי אדמה מועכים אותם ברגל גאוה [על פי תהלים לו, יב]. מאז קוראים בביתי לממעך רגל גאוה, אבל כמובן אינני מציעה שהאקדמיה תאמץ זאת. אני מציעה לקרוא לממעך תפוחי האדמה פִּיֶּרֶן – על שום הפִּיֶּרָה – גם אם יש בכך מתן הכשר למילה לועזית שנקבעה לה חלופה עברית.

א' ממן: אני בעד ממעך ובעד הצעתו של האד' בלאו לבוידם עלייה. ליודעי לשונות

שמיית אחרות יכולה להיות בעיה עם המילה אחסן – הן בשל מובנה השונה של אחסן (במלעיל) בערבית הן בשל הזיקה שבצליל אל אחסנתא בארמית.

ד' טלשיר: גם דעתי אינה נוחה מאחסן, ואני מציע תחתייה עליית מחסן. אשר לממעך, אינני רואה במה הוא טוב ממה שנקבע במילוני האקדמיה – מְרִסָּק.

ק' כ"ץ: מבחינה מעשית לא נכון לקרוא לשני כלים שונים המשמשים באותו תחום בשם אחד. אם יבקש אדם מרעהו להביא לו ממעך מן המטבח, לא ידע השומע מה מבוקשו.

א' טל: במה שנוגע לבוידם אינני חושב שיש צורך בחידוש ואפשר לקרוא לו עלייה. אין העלייה של היום דומה לעליות שנבנו בעבר, ולכן אין לחשוש מבלבול. כותש השום באמת כותש את השום, לא מועך אותו. לפיכך השם העממי מתאים, משמש ואינו מתנגש עם מונח אחר, ואני מציע להניח לו כפי שהוא.

מ' בריאשר: אם אין דוברים נוספים, נעבור להצבעות. לבוידם שמענו כמה וכמה הצעות, חלקן חדשות. כדי למנוע חשש שתתקבלנה החלטות כלאחר יד, נצביע רק על אלה שהיו לפני הוועדה. אפשרות אחרת היא להחזיר את המונח לוועדה לדיון מחודש כדי לשקול את ההצעות החדשות שעלו כאן.

הצבעה

| | |
|-----|------------------------|
| 1 – | בעד אַחְסָן |
| 9 – | בעד עֲלִיָּה |
| 1 – | בעד עֲלִיַּת תְּקָרָה |
| 9 – | בעד החזרת המונח לוועדה |

מ' בריאשר: נצביע שנית כדי להכריע בין שתי ההצעות שקיבלו את מספר הקולות הרב ביותר.

הצבעה

| | |
|------|------------------------|
| 11 – | בעד עֲלִיָּה |
| 9 – | בעד החזרת המונח לוועדה |

הוחלט: עֲלִיָּה (בוידם).

[הערת העורך: בישיבה רנט (להלן, עמ' 187) תוקנה ההחלטה לעֲלִיַּת תְּקָרָה (ובקיצור עֲלִיָּה)].

מ' בריאשר: לא שמעתי התנגדות למונח המוצע אזור תפעול. נצביע עליו, ואחר כך נכריע בין ממעך – הן לתפוחי אדמה הן לשום ולשאר ירקות – ובין ממעך לתפוחי אדמה וכותש שום לשום כמובן.

הצבעה

בעד אזור תפעול – הכול

הוחלט: אזור תפעול (ground support area).

11 – בעד מְמַעֵף (לכול)

13 – בעד מְמַעֵף (לתפוחי אדמה) ולצדו פּוֹתֵשׁ שוּם

הוחלט: מְמַעֵף (potato masher); פּוֹתֵשׁ שוּם (garlic press).

מ' בר-אשר: נפנה עתה לחמש המילים שנותרו, ובראשן חלופה עברית לפרמדיק. אני מבין שהיא מהווה בעיה מיוחדת.

מ' קיסטר: מצד הלשון ההצעה רפאי חירום חלף פרמדיק טובה. גם הרופאים שפנינו אליהם לא הביעו הסתייגות. ואולם התנגדות נחרצת באה מכיוונו של ציבור היעד – הפרמדיקים. למן הישיבה הראשונה התנגד יושב ראש ארגון הפרמדיקים לעצם הרעיון של החלפת שמם, בלי קשר לחלופה העברית שהוצעה. הוא דיבר קשות ואף אמר שאם ראש מגמת "רפואת חירום" בבאר-שבע יאמץ את החלטת האקדמיה, הוא יביא לסגירת המגמה.

אקרא את עיקרי הנימוקים ששלח לנו בתגובה דובר "מגן דוד אדום" ושמהם מתחזרת חומרת ההתנגדות – לא התנגדות עניינית אלא עקרונית: "בבחירת השם רפאי ישנו חשש אמתי וסביר שהפרמדיק ייתפס בקרב הציבור כעוד מקצוע שרלטני בקרב המרפאים. אנו מתקשים להבין מדוע אתם נחפזים לבחור שם אחר לפרמדיק, שהפך כבר מזמן למטבע לשון שכל בר בי רב מכיר ומבין. בדרך כלל כאשר נמסר למשפחה שקוראת לאמבולנס שהפרמדיק בדרך, יש למשפחה הרגשת הקלה ותחושת ביטחון שמגיע איש צוות רפואי בעל הכשרה מקצועית. השימוש במילה בין-לאומית מקל על התקשורת הענפה עם ארגונים דומים בעולם, בדומה לתואר אקדמיה שנושא מוסדכם. גורלה של מילה שתקבע האקדמיה במקום פרמדיק יהיה כגורל מילים אחרות שלא נקלטו ושזוכות לעדנה רק במשחקי טריוויה".

א' ממנן: ההיגיון אומר שרפאי חירום הוא סוג של רפאי. ואולם ההצעה רפאי נדחתה, וכך יש כאן צירוף שאחד מרכיביו אינו עומד לעצמו.

ר' גדיש: אחד הנימוקים להוספת הרכיב חירום היה קרבת השם רפאי לרְפָא, מי שעיסוקו בשחזור כלי חרס.

מ' בר-אשר: כפי שאני רואה את הדברים השאלה היא אם לקבוע מונח שצרכנו דוחים אותו מראש. אני חושב שרצוי להסיר את המילה מסדר היום.

י' בלאו: אף שרפאי ביחס לרופא נשמע כהנדסאי ביחס למהנדס, יש לשים לב שלא הרי היו"ד של הנדסאי, שהיא חלק של הסיומת, כיו"ד של רפאי, שהיא מן השורש. ואולם נראה לי שההתנגדות לרפאי חירום נובעת מחשש לפגיעה כלשהי בערכו של המקצוע. אני מסכים עם ההצעה להסיר את המילה מסדר היום.

ההצעה רפאי חירום (פרמדיק) הוסרה מסדר היום.

מ' קיסטר: אף שלא ידוע לנו בעברית על מילה לצהלולים, ברור שמצהלות השמחה האלה נהגו בעם. אני מכיר לכך התייחסות במגילות קומראן. במגילת "מלחמת בני אור בבני חושך" מצינו "בנות עמי, הפענה בקול רנה", ובמקום אחר במגילה כתוב "צרחנה בקול רנה".

מ' עסיס: השם מקצוע נלווה רפואה נראה לי מסורבל מדי. לדעתי דרושה כאן הצעה פשוטה יותר.

המילה צמודונים יפה, אבל אינני מבין את פשר הסיומת ין, המשמשת פעמים הרבה להקטנה.

נ' שמר: גם בגרבוך אין לסיומת זו משמעות של הקטנה. הצמודונים באים בעקבות הגרבוכים.

י' בלאו: בעניין מקלור אני תוהה אם כדאי לקבל מילה חדשה, או שמא עדיף רק לקבוע צירוף חדש שרכיביו ישנים – מקל-אור. כך גם נפתרת הבעיה אם לקיים את הדגש בקו"ף השוואה או לא.

ר' גדיש: את ההצעה מקלור כבר מצא דורון רובינשטיין, עובד המילון ההיסטורי, בספר הלכות חדש.

נ' שמר: גם בעיניי מקצוע נלווה רפואה נראה מסורבל. אני רוצה להעלות הצעה למחשבה שתישמע תחילה מוזרה במקצת: להשתמש באות היחס כ"ף ולקרוא למקצועות האלה מקצועות פּרפּוּאָה.

א' ממך: לאזנני הצירוף נלווה רפואה נשמע מצוין. אשר לצהלולים, אני מכיר את השימוש בפועל יליל. הוא בא גם במילון אבן-שושן, ולכן אינני חושב שהחידוש נחוץ.

מ' בראשור: במקומות שניסיתי להפיץ את דבר הצהלולים, נתקבל החידוש בשמחה.

י' בן-טולילה: אינני חושב שהרכיב אור מתאים למקלור. המקלור אינו מאיר אלא זוהר, ועל כן הייתי מחפש מילה מן השורש זה"ר. למשל זְהָרוּר על משקל גפרור. גם זוהר וגם מזהיר.

מ' בר-אשר: מאחר שזאת הצעה חדשה, נעביר אותה לדיון בוועדה למילים בשימוש כללי אם ההצעות מקלור או מקל-אור לא תתקבלנה.

הצבעה

בעד (מקצוע) נְלוּה רְפוּאָה – 15

הוחלט: (מקצוע) נְלוּה רְפוּאָה (מקצוע) פרה-רפואי, (allied health professional).

מ' בר-אשר: ל-stick-light נשמעו שלוש הצעות: מקלור, מקל-אור וזהרור. מאחר שההצעה וזהרור חדשה, נעביר אותה לדיון בוועדה למילים בשימוש כללי אם שתי ההצעות הראשונות לא תתקבלנה.

הצבעה

בעד מְקָלוֹר – 14

בעד מְקָל-אוֹר – 2

הוחלט: מְקָלוֹר (stick-light).

בעד צְהָלוּלִים – רוב גדול

הוחלט: צְהָלוּלִים (زغردة, ברבים: زغاريد; קולות שמחה המקובלים בתרבות המזרח).

בעד צְמוּדוֹנִים – הכול

הוחלט: צְמוּדוֹנִים ("טייטס", מכנסיים צמודים; (tights, [panty] hose).

ג. הוועדה למינוח טכנולוגי בחיפה

מ' בר-אשר: הסעיף האחרון שעל סדר היום הוא הלשכה למינוח טכנולוגי בחיפה. תחילה אבקש ממנכ"ל האקדמיה להביא לפני האקדמיה את הודעתנו בעניין זה.

נ' אפרתי: בחודשים האחרונים חלו התפתחויות רבות בנושא לשכת חיפה, וכדי להביא לפני חברי האקדמיה את הדברים כהווייתם, אקרא לכם ביטויים מכמה מסמכים; הראשון הוא מכתב נשיא הטכניון עמוס לפידות מד' באלול התש"ס (4 בספטמבר 2000). אקרא את תחילתו:

מזה עשרות שנים מתקיימת בטכניון פעילות חשובה ביותר להעשרת השפה העברית במונחים חדשים בתחומי הטכנולוגיה. פעילות זו מתבצעת באמצעות

הוועדה המרכזית הפועלת בטכניון, והיא נגזרת מתחומי האחריות של האקדמיה ללשון העברית ובשיתוף פעולה עמה. בשנים האחרונות חלו תקלות והפרעות בפעילות הוועדה אשר התבטאו בשיבושים בקשר ובשיתוף הפעולה עם האקדמיה ללשון העברית, בהפסקת תקצוב עבודתו של מזכיר הוועדה והתרחקותם של פעילים בכירים, למשל הפרופסורים אלדר ואזר, ובאי-בהירות לגבי סמכות, אחריות וסדרי מנהל. לאור הדברים הללו מינה נשיא הטכניון לפני שנתיים ועדה מטעמו לבדוק את הדברים.

דוח הוועדה הוגש לנשיא הטכניון בתחילת ספטמבר 1999. דוח משלים הוגש לו באמצע פברואר 2000. ב-3 ביולי השתא התקיימה פגישה בין נשיא הטכניון לנשיא האקדמיה, ולבסוף הגיע נשיא הטכניון למסקנות האלה (אני קורא מן המסמך שלו):

הגם שנושא המינוח הוא בתחום האחריות הממלכתית של האקדמיה ללשון העברית, ימשיך הטכניון לעסוק בכך בהעריכו את החשיבות הלאומית... הפעילות תיעשה בשיתוף פעולה מלא עם האקדמיה ללשון העברית ובכפוף לכללי האקדמיה לגבי פעולתן של ועדות מינוח בתחומים השונים. עם חידוש פעילות המינוח הטכנולוגי, יועסקו שני נושאי משרה בשכר, יוצאו מינויים חדשים לחברי הוועדה המרכזת למונחי טכנולוגיה וירועננו סדרי העבודה של הוועדה. הסמכויות והמשימות בתחום המינוח הטכנולוגי תהיינה דומות לאלה של ועדות האקדמיה בתחומי מינוח אחרים, ולאור זאת ההחלטות טעונות אישור ועד המינוח של האקדמיה.

נשיא הטכניון החליט אפוא לארגן מחדש את הוועדה המרכזית למונחי הטכנולוגיה ולחדש את פעולתה בצורה מקפת ובשיתוף פעולה מלא עם האקדמיה. המכתב הופץ לידיעתם של כל הנוגעים בדבר, ואולם כעשרה ימים לאחר מכן שלח האד' אירמאי מכתב לכל שרי ממשלת ישראל ונושאי משרות בכירים אחרים, כגון מבקר המדינה והרמטכ"ל, בנושא "סגירת הלשכה למינוח טכנולוגי בטכניון לאלתר". בין היתר כתב האד' אירמאי כך:

ללא כל התראה מוקדמת החליט הטכניון לסגור את לשכת הוועדה... הוועדה היא גוף סטטוטורי הפועל מכוח חוק המוסד העליון ללשון העברית... התוצאה תהיה משבר מינוח לזמן ממושך, תופסק הפעילות המשותפת והבין-מוסדית לתאום המינוח בין כל הגורמים המעורבים בהיערכות העורף והמשק לחירום, ולא יוכל להינתן עוד שירות מינוח טכני מוסמך ללשכה המרכזית לסטטיסטיקה, למשרד ראש הממשלה, למבקר המדינה וכיו"ב.

בעקבות מכתבו של האד' אירמאי קיבל נשיא האקדמיה מכתב משר המדע, התרבות והספורט, שזנעק לעניין המינוח הטכנולוגי וחרד לגורלו. ברשותכם אקרא כמה ראשי פרקים מתשובתו של נשיא האקדמיה לשר מתן וילנאי:

ראש לכול, כדי להסיר דאגה מלבך יכול אני להרגיעך כי הלשכה למינוח

טכנולוגי לא תיסגר. ועתה ברשותך אפרט בקצרה את השתלשלות הדברים. הצד החוקי: סעיף 10 לחוק המוסד העליון ללשון העברית מוסר בידי האקדמיה את הסמכות לקבוע החלטות בענייני דקדוק, כתיב, מינוח או תעתיק. הוועדה המרכזית למונחי הטכנולוגיה, המשותפת לאקדמיה ולטכניון, הוקמה על פי החלטת מליאת האקדמיה בישיבתה מ"ג בתמוז תשט"ו - 1956. לוועדה אין אפוא כל מעמד סטטוטורי, לא מכוח חוק ולא מכוח תקנון, וכינונה מסור אך ורק בידי האקדמיה.

אני מבקש להוסיף ולומר כי באותה ישיבה שבה הוחלט על הקמת לשכת חיפה אמנם השתמשו במונח תקנון - בזיכרון הדברים נכתב שאושר תקנון הוועדה המרכזית - אבל בסופו של דבר אין מדובר אלא בהסכם בין הטכניון לאקדמיה. ברצותם שני השותפים שומרים על ההסכם, וברצותם הם מביאים אותו לכלל סיום. ההסכם לא עבר הליך רשמי של פרסום ברשומות. אין לו משמעות משפטית מעבר למה שהצדדים נתנו לו מלכתחילה, וחשוב שהדברים האלה - המנוגדים לפרשנויות שיוחסו להסכם עם השנים - ייאמרו כאן במליאה.

עוד כתב נשיא האקדמיה בתשובתו לשר כך:

האקדמיה מפעילה כל העת ועדות מינוח בתחומים טכנולוגיים, הן בשיתוף הטכניון הן בשיתוף מכון התקנים. עמוס לפידות, נשיא הטכניון, קבע כי בשנים האחרונות חלו תקלות והפרעות בפעילות הוועדה אשר התבטאו בקשר ובשיתוף הפעולה עם האקדמיה. בשנה וחצי האחרונות פעלה בטכניון ועדה שמינה נשיא הטכניון לבחינת הנושא, והמלצותיה נדונו בפורום של חבר הנהלת הטכניון בראשות הנשיא. ההחלטה המרכזית שהתקבלה בהנהלת הטכניון הייתה לחדש את פעילותה של הוועדה המרכזית למונחי טכנולוגיה, פעילות שתיעשה בשיתוף פעולה מלא עם האקדמיה ובכפוף לכללי האקדמיה לגבי פעולתן של ועדות מינוח בתחומים השונים. לאור זאת מינתי את פרופ' אור ליושב ראש הלשכה למינוח טכנולוגי.

לסיכום, עבודת המינוח הטכנולוגי שפרופ' אירמאי חרד לה נמשכת כל העת, והיא תתחדש עתה ביתר שאת בצורה מסודרת בעקבות התיאום וההסכמה המלאה שבין האקדמיה והטכניון, על פי כללי המנהל התקין.

אני מצר מאוד שפרופ' אירמאי, הנוכח כאן, שצבר כל כך הרבה זכויות במשך עשרות שנים בתחום המינוח הטכנולוגי, הגיע למצב שבו לא החמיץ שום במה פומבית לתקוף את האקדמיה. לדוגמה אקרא לכם כמה שורות מכתבה שהתפרסמה בעיתון "הארץ" ב-30 בנובמבר 2000:

המבט בעיניו של פרופ' אירמאי מעיד שהוא יודע שמפעל חייו הגיע לנקודה סופית. הלשכה למינוח טכנולוגי בטכניון, שמנותקת מתקציבים במשך ארבע שנים, מפסיקה מחר לפעול באופן רשמי. בין אם תיוולד מחדש תחת שם אחר ובין אם תנדוד אל משכן אויבתה הגדולה, האקדמיה ללשון העברית.

וכן:

הצרות החלו כמעט מיד. האקדמיה לא אהבה את הגוף המקצועני והמרוחק מחיפה שבו לא שלטה ישירות.

ובסוף:

השנה החליט נשיא הטכניון עמוס לפידות לשים קץ לסיפור. הוא הודיע על גגירת הלשכה בטכניון והוציא מכרז חדש למזכיר הלשכה.

אם כן, חשוב שמליאת האקדמיה תשמע על ההתפתחויות האלה ושתדע שפעילות הוועדה למינוח טכנולוגי נמשכת כל העת בצורה זו או אחרת. למרות הפרסומים השונים החליטו האקדמיה והטכניון להמשיך את פעילות הלשכה, בתקווה לייצל את עבודתה ולשפרה.

מ' בראש: תודה לאד' אפרתי על הצגת הדברים. אני מעוניין להדגיש ולומר כי חרף דברי ההשמצה והגידופים כלפי האקדמיה אסור שהאקדמיה תתעלם לרגע מזכויותיו רבות השנים של האד' אירמאי. ואולם כל עניין לשכת חיפה בא מחיפה. ראשי הטכניון פנו אליי בשנים האחרונות בתביעות חוזרות ונשנות לסגור את הלשכה. לא קיבלנו תביעה זו. אבל מאחר שחברי האקדמיה שהיו פעילים בעבר בלשכה הדירו רגליהם ממנה בטענה שאינם זוכים לשיתוף פעולה, כשבא נשיא הטכניון וביקש עזרה בסגירת הלשכה ובפתיחתה מחדש, הבנו שאין מנוס מנקיטת צעדים מסוימים.

ש' אירמאי: כבוד הנשיא והחברים, לפני מספר שנים, כשביקשתי לספר כאן במליאה על הלשכה, לא ניתן לי להשלים את דבריי. זאת הייתה הפעם היחידה שהמליאה קיימה דיון על לשכת חיפה.

אני מצר שהאד' בן-חיים איננו כאן. הוא יכול להעיד על מעמדו של ההסכם בין האקדמיה לטכניון ולחזק את דבריי שמדובר בתקנון רשמי. היו אלה ימים של ראשית המדינה, וגם האקדמיה לא נהגה לפי הכללים ולפי המתחייב ממנה בחוק. למשל, חוק המוסד העליון ללשון העברית קובע מה צריך תקנון האקדמיה לכלול. קרוב לחמישים שנה עברו, ועדיין לא באו הדברים על תיקונם. בשנים 1976-1980 הייתי חבר במנהלת האקדמיה, ולבקשת הנשיא בן-חיים הוטל עליי להכין טיוטה מפורטת ומשלימה לתקנון האקדמיה וועדותיה. לצערי הרב ההצעות שהגשתי לאקדמיה נשארו חתומות בתיק כאבן שאין לה הופכין, והוועדה המרכזית נותרה עם מסמך יסודה משנת 1955.

הוועדה המרכזית פרסמה במהלך למעלה מחמישים שנותיה כחמישים מילונים טכנולוגיים. היא העסיקה ארבעה עובדים; שניים מהם על ידי האקדמיה. אני שימשתי מזכיר הוועדה למן 1943 ויושב ראש הוועדה בשנים 1969-1993, מה שיצר קשר אמיץ בין הטכניון לאקדמיה. עם בחירת בני רון למזכיר הוועדה התפטרתי מתפקידי כיושב הראש שלה. במקומי בא האד' שמעוני. עם צאתו לשבתון החליפו למשך שנה האד' אור, ועם שובו מן השבתון המשיך האד' שמעוני למלא את תפקידו.

בשל נימוקים תקציביים הופחת בהדרגה בשנים האחרונות – תחילה באקדמיה ואחר

כך גם בטכניון – מספר העובדים. בתחילה לעובד יחיד (המזכיר), אחר כך לחצי משרה, ובשנת 1998 נפסק התקציב כליל והמזכיר רון אירמאי פוטר. או אז ביקש ממנו היו"ר בפועל האד' שמעוני להמשיך בתפקידו בהתנדבות עד שהטכניון יחליט סופית על עתיד הוועדה. רון הסכים והמשיך בתפקידו זה שנתיים בלי תקציב ומשכורת. כיוון שהמעוניינים בביטול הוועדה המרכזית בטכניון לא יכלו לבוא בטענות אליי, הם הפנו אותן נגד רון אירמאי המתנדב.

מ' בר-אשר: תודה לאד' אירמאי, אני נועל בזה את הישיבה.