

מִשְׁאַת אֶהֱרֵן

מחקרים בלשון מוגשים לאהרן דותן

עורכים

משה בר־אשר, חיים א' כהן



מוסד ביאליק • ירושלים

תוכן העניינים

ה		פתח דבר
1	רשימת הפרסומים של פרופ' אהרן דותן	נאסר בסל
		לשון המקרא ולשונות שמיות
13	על שני שוואים נעים רצופים פרשנות ימי הביניים על ספר בראשית לאור הפילולוגיה המקראית בת זמננו – חלק שני:	יהושע בלאו חיים כהן
25	בראשית יט-לה הסביל הפנימי של הבניין הקל בלשונות	הניאל סיון
47	השמיות המערביות	
57	סופית הייחוס יון בעברית הקדומה	אלישע קימרון
		לשון חז"ל, לשון התפילה וארמית
	ניקוד הלמ"ד במילת של- במשנה במסורתו	גבריאל בירנבאום
63	של מנקד כ"י קאופמן	
75	מקצת הַרְבָּה	ישראל בן-דוד
92	משקל פְּעֻלַת בלשון המשנה לשון תפילת העמידה של חול לפי הנוסח שבסדר	משה בראש-ר יצחק גלוסקא
102	רב סעדיה גאון מסורות קריאתן של מילים מקראיות ובתר-מקראיות	עמוס דודי
126	במחזורי קטלוניה	
139	קטעי תרגום אסתר מן הגניזה והעדויות שהם מעידים תצורת הבינוני בבניינים הסבילים פועל והופעל	אברהם טל חיים א' כהן
172	בלשון המשנה על פי כתב יד קאופמן A50	
189	עוד תפילה של השומרונים אב חסדה הצורי	משה פלורנטין
196	גינני גינייה, גנתה גננה, גנונית גנונייה קטע גניזה של ברכת המזון: מחקר בנוסח,	מרדכי עקיבא פרידמן שלמה קלמן רייף
201	במחשבה ובתולדות התפילה	
218	חילופי העיצורים ב-פ בלשון חכמים	שמעון שרביט

	נאסר בסל	על הלוואי בעיבוד בערבית-יהודית לחיבורו של
247		אבן ג'ני "אללמע פי אלערכיה"
	דן בקר	"האותיות הנוספות" – משמען אצל המדקקים
		העברים בימי הביניים לעומת משמען אצל עמיתיהם
261		הערבים
280	מרדכי ברויאר (ז"ל)	שווא נע ושווא נח על פי כללי הטעמים
	רעיה חזון	ספר תיקון השגגות – תשובות אדניה על רס"ג: זהות
289		מחבר התשובות וסוגיות נוספות
	רחל חיטין-משיח	"אהוב לבי השיבני תשובה": עוד פתרון לשיר חידה
305		תימני על טעמי המקרא
316	יוסף טובי	יסודות מילוניים בפירושי רב סעדיה גאון למקרא
358	אהרן ממן	קטעים חדשים מכתאב אלמואזנה לר' יצחק אבן ברון
392	יוסף עופר	רשימה בכלית של פרשיות פתוחות וסתומות בתורה
435	מאיה פרוכטמן	המדקק כמשורר והמשורר כמדקק בימי הביניים
	מיכאל ריז'יק	הערות ללשון המכתבים העבריים מטוסקנה
442		מן המאה החמש-עשרה
	נורית רייך	ה"זברקונים" שבברכת יעקב לבניו: מדרש לשון
453		והערת מסורה לפתרון צימוד ב"כתאב אלתג'ניס"
469	רונית שושני	תפקידם המקורי של טעמי המקרא
	ראובן ש' שטיינר	פתח וקמץ – עיון בגיזרוןם ובהתפתחותם של
487		שמות התנועות בעברית
498	אליעזר שלוסברג	פירושי רס"ג במדרשי תימן המאוחרים
	העברית החדשה	
525	אילן אלדר	החייאת הדיבור העברי
	אסתר בורוכובסקי	הגדרה מילונית למילים מחלקי דיבר שונים – גישות
545	ופנינה טרומר	שונות אך משתלבות
	דויד דורון	עיוני תחביר בעברית שבפרסומים של הרשות
565		הפלטטינית
590	שלמה יזרעאל	מדיבור לתחביר – מתאוריה לתמליל
	יצחק צדקה	מ"ענדו כתאב" (אצלו ספר') ל"ענדויאה"
611		(אצלו אותר')
	אורה (רודריג) שורצולד	פנים נחוץ בעברית: בין התפתחויות היסטוריות
619		טבעיות להשפעות זרות
630	יצחק שלזינגר	תפוצת שמות התואר בשיח השכנוע

החלק הלועזי

3*	הגייתו של שווא-געיה על פי המסורת הטברנית	ג'פרי כאן
	בין אברהם אבן עזרא לרב סעדיה גאון בשאלות	אנחל סאנז-בדיליוס
19*	של דקדוק	
33*	האותיות הדבוקות במסורה ובהלכה	מישל צרפתי
46*	תרגום יחזקאל על פי המסורת הבבלית	יוסף ריברה-פלוריט (ז"ל)

מדיבור לתחביר – מתאוריה לתמליל

שלמה יזרעאל

הקדמה: מאגר העברית המדוברת בישראל (מעמ"ד)

במאמר זה אביא בפני הקורא הרהורים אחדים על דרכי רישום הדיבור לצורכי מחקר. מכיוון שמחקר הלשון הדבורה¹ עודנו בחיתוליו (ראה מילר וויינרט, 1998, סעיף 1.2; בייבר ועמיתיו, 1999, פרק 14),² וחקר העברית המדוברת עודנו מצפה לביסוס ראשיתי (יציב-מליבר, 2002, פרק 1; יזרעאל, תשס"ב, א), אין בדברים דלהלן אלא שאלות וכיווני חשיבה שעלו תוך כדי קריאה ועבודה.

עבודתי בתחום זה נבעה מן הצורך לגבש כללים מנחים להעברת הטקסטים העתידיים להיכלל במאגר העברית המדוברת בישראל (מעמ"ד) אל ערוץ הכתב. מעמ"ד יכונן כמאגר מייצג. כדי להבטיח ייצוג אמין של השונות הלשונית, הן זו הדמוגרפית והן זו התלויה בנסיבות השיח, יישמר עקרון האקראיות בעת דגימת הנתונים הטקסטואליים. דגימת נסיבות השיח תתבצע לאחר איסוף כל ההקלטות של האוכלוסייה הנדגמת דגימה דמוגרפית. כל אלה שעלו במדגם האוכלוסין האקראי יתבקשו להקליט את כל הקורות אותם במשך יממה רצופה. כל הקלטה כזאת תיבדק, ויוסרו ממנה מרווחי שתיקה וקטעי דיבור לא ברורים. מן החומר הנקי ייבחר קטע של טקסט באורך של כחצי שעה עד שעה (בערך 5,000 מילים), והוא שיהיה מקור הנתונים שייכלל במאגר הראשי. לפרטים על תכנון המאגר ועל תכנית העבודה ראה יזרעאל, הרי ורהב, תשס"ב; הרי וזרעאל, 2003. על מצאי ההקלטות הקיים ראה יזרעאל, תשס"ג וזרעאל ורהב, 2004.

- * תודה לנועם אמיר, ליעקב בן-טולילה ובמיוחד לורד זילבר-וורוד על הערותיהם החשובות.
1. הצירוף לשון דבורה מכוון אל מרחב חלופות השפה הספונטנית והטבעית. הצירוף לשון מדוברת משמש לסימון ערוץ התקשורת שבדיבור על כל חלופותיו, ועומד לעומת ערוץ הלשון הכתובה על כל חלופותיו. ראה יזרעאל, תשס"ב, א, עמ' טו-טז.
 2. אמנם הפונטיקה, הפונולוגיה וההנגנה – שהיא עיקר מושא העיון דלהלן – נחקרו לא מעט, אבל תחומים אלה – לצד שאר תחומי המחקר הבלשני – לא נחקרו כהלכה בלשון הדבורה על גווניה השונים.

תמליל ותעתיק

“תעתיק הוא תאוריה” (Transcription is theory). במילים אלה, המבוססות על כותרת מאמר חשוב מאת אלינור אוקס, “תעתיק כתאוריה” (אוקס, 1979), פותח ג'ון דו בואה את מאמרו הדין בעקרונות העברת שיח מדובר לערוץ הכתב למטרות מחקר (דו בואה, 1991). משוואה זו היא הבסיס שאין בלתו לחקר הלשון המדוברת. transcription מתורגם ברגיל תעתיק, אולם במאמרים העוסקים בנושא זה, ובהם מאמרו של דו בואה, הרישום הנוהג אינו תעתיק פונטי אלא העברה אל הכתב בכתיב המקובל, בשינויים קלים מתבקשים ובתוספת סימנים לרישום יסודות פרוזודיים או לרישום יסודות לא מילוליים. להלן ישמש המונח תמליל לציון העברה אל הכתב בכתיב המקובל,³ והמונח תעתיק ישמש לציון רישום הדיבור בסימנים פונטיים (באותיות לטיניות או בסימנים מקובלים אחרים, כגון סימני התעתיק השמי או סימני IPA).

תמליל, כמוהו כתעתיק, מעוגן במסד תאורטי. דרכי התמלול והתעתוק – ותוצריהם, התמליל והתעתיק – תלויים במטרות המחקר. עניין מהותי זה נרמז בכותרת מאמרו של דו בואה, חבר צוות מעמ"ד וראש מאגר האנגלית האמריקאית המדוברת של סנטה ברברה (Santa Barbara Corpus of Spoken American English),⁴ שעבודותיו משמשות בסיס לעבודת התמלול של מעמ"ד (וראה עוד להלן).

ממאפייני המאגר, מהיקפו וממטרותיו נגזר אופייה של דרך הצגת הטקסטים. מעמ"ד יקיף, על פי הגדרתו, מגוון רחב של נושאי מחקר, ובהם בלשנות, חינוך, היסטוריה וזהות לאומית ותרבותית וכן הנדסת לשון. צוות המאגר ראה לנכון לא להגביל את שימושי המאגר ליישומים מוגדרים, אלא לאפשר מחקר רבגוני. מאגר בהיקף גדול מכתוב את הצורך לחסוך עד כמה שאפשר בזמן עבודה, בכוח אדם ובהוצאות. הטקסטים שייכללו במעמ"ד יהיו טקסטים שהוקלטו בסביבה טבעית, רועשת לעתים, ובנסיבות טבעיות. מכאן שהעברת הטקסטים לערוץ החזותי אינה דבר של מה בכך. ניסיון אחרים הידוע לנו הוא ניסיונם של מכונני מאגרים של לשונות הודו-אירופיות, במיוחד אנגלית. המאגרים הקיימים אינם מתועתקים אלא מתומללים, כלומר שיטת העברת הערוץ הנקוטה בידם משתמשת בכתיב הנהוג בתוספת סימנים מוסכמים. מניסיון אחרים ומניסיון שצברנו אנו נדמה כי תמלול איכותי של הדיבור עשוי להצריך בין 100 ל-300 שעות לכל שעת הקלטה. מן האמור לעיל נובעת, קודם כול, ההחלטה שהקלטות מעמ"ד לא יוגשו לקהל היעד בתעתיק פונטי, הדורש שעות עבודה רבות פי כמה וכמה מהשעות הנחוצות

3. האקדמיה ללשון העברית טבעה את המונח תכתוב להעברת ערוץ ככלים אלה. גם שם הפעולה יכונה תכתוב. אני מעדיף לדבוק במונח תמליל ולהכביל ממנו את שם הפעולה תמלול. שורש המונחים תמליל ותמלול משקף טוב יותר, לדעתי, את העברת רצף הדיבור ואת חלוקתו למילים. ועוד: המונח תכתוב אינו נפרד דיו מן המונח תעתיק; שניהם אפשר שיסמנו העברת ערוץ, אולם לא את מהות ההעברה. ולבסוף: המונח תכתוב קרוב מדי לשם תכתוב.

4. ראה <http://www.linguistics.ucsb.edu/research/sbcorpus.html>; <http://projects ldc.upenn.edu/SBCSAE/>.

לתמלול; מה גם שמעמ"ד יעמיד את הקלטות הדיבור עצמן לרשות המשתמשים. אולם ההחלטה הזאת נובעת לא רק משיקולים מעשיים אלא גם משיקולים של תאוריה. לשאלת הגשת הטקסט המדובר בתמליל ולא בתעתיק ראה יזרעאל, 2004; והשווה יזרעאל, תשס"ג-תשס"ד.

הלשון המדוברת ותיאורה

כותרת מאמר זה, "מדיבור לתחביר – מתאוריה לתמליל", נראית כיאסטית לכאורה, ואולי אף כאוטית. אפשר היה לצפות לכותרת מעין "מדיבור לתמליל – מתאוריה לתחביר" או שמא "מדיבור לתאוריה – מתמליל לתחביר", ואולי "מדיבור לתאוריה – מתחביר לתמליל". הכותרת כפי שהיא הריהי כמובן מכוונת, ומכוונת לדעה שללשון המדוברת יש תחביר משלה, ושהתאוריה שעל פיה נתאר את תחביר הלשון המדוברת היא שתוביל לעיצוב תמליל הטקסטים בלשון המדוברת.

מבנה הלשון המדוברת שונה ממבנה הלשון הכתובה. אם נתבונן בלשון המדוברת כישות מבנית משל עצמה, ונתעלם לרגע מקשריה עם הלשון הכתובה, נראה שאפשר לכתוב דקדוק כוללני ללשון המדוברת שיהיה שונה במהותו מן הדקדוק שאנחנו רגילים לראות. לא רק התחביר יהיה שונה, אלא אף מבנה השיח, אוצר המילים, הסמנטיקה הלקסיקלית והסמנטיקה של הדקדוק, ואפילו המורפולוגיה והמורפופונולוגיה.

לא הזכרתי את הפונולוגיה, ולא בכדי. לאמתו של דבר, המונח פונולוגיה מטעה אם מכוונים אנו ללשון הכתובה בלבד, כי אין ללשון הכתובה צלילים. אם המונח פונמה יוגדר כיחידת הלשון שמרכיביה צלילים (השווה למשל מתיוס, 1997, עמ' 276) ומערכת פונמית תוגדר כמערכת הגאי הלשון (שם, עמ' 278), הרי ללשון הכתובה אין פונמות; יש לה אותיות, יש גרפמות. גם אם נקרא בקול מן הכתב, העברת הערוץ תתבצע תוך כדי תרגום של היחידות המינימליות מן הלשון הכתובה אל הלשון המדוברת. למשל, לשון הכתב בעברית תבחין בין ט לבין ת, אולם לא לשון הדיבור.⁵ לעומת זאת, לשון הדיבור תבחין בין p לבין f, אולם לשון הכתב – הכתב הבלתי-מנוקד, כמובן, שהוא דרך הכתיבה והקריאה הרגילה לנו – לא תבחין בין העיצורים הללו (ראה עוד יזרעאל, 2004; יזרעאל, תשס"ג-תשס"ד).

בחברה אוריינית כמו החברה שלנו – שלה מסורת ארוכה של קריאה וכתיבה וגם מחקר דקדוקי בן אלף שנים ויותר – ההישענות המטה-לשונית על הלשון הכתובה רבה ועמוקת שורשים. בדברנו על מרכיבים של לשון הדיבור אנו מדברים בעצם על מרכיבים של הלשון הכתובה. איננו מדברים על קולות או על פונמות כי אם על אותיות; אנו חושבים שאנו מדברים במשפטים ומכנים את יחידות התחביר של הלשון המדוברת

5. לבד מבפיהם של זקנים בני עדות אחדות.

במונח זה; ואנו משתמשים במושגים מלשון הכתב בדברנו לא רק על הלשון אלא גם על נושאים אחרים (ראה בורוכובסקי-בראבא, תשס"ב).

המודל הדקדוקי שאנו רגילים לבנות ללשון המדוברת נשען על מודלים שנבנו ללשון הכתב. לא רק הבלשנות השמית אלא הבלשנות בכלל נגועה במסורת רבת דורות זו, שכן מקובל להניח שחקר הלשון יכול להתבצע אך ורק על סמך נתונים כתובים. אני לא בדקתי את ההנחה הזו, אולם חקר תולדות הבלשנות תומך בה תמיכה ממשית (וראה גם אהליך, 2005, עמ' 104-106). והרי גם את מבנה הלשון המדוברת אנו בודקים רק לאחר שיש בידינו העברה של הלשון המדוברת אל הערוץ החזותי.⁶

תמליל וסימוניו

תצוגת הטקסטים של מעמ"ד לקהל המשתמשים תשאף להביא את הדיבור המוקלט מלווה בתמליל סינכרוני, שבו ההקלטות ותמליליהן מוצגים במקביל: התמליל מופיע בד בבד עם רצף הדיבור. שיטה זו מאפשרת למשתמש גישה בלתי אמצעית לכל מאפייני הדיבור. עם זאת דרך זו לא תאפשר סקירת רפרוף של הטקסטים, שהרי אין בה גישה לטקסט הרציף, וגם לא שימוש בחומרי המאגר לשם פרסום מחקרים. כדי לאפשר שימוש בתמליל גם בלא ליזוי הטקסט הקולי, התמליל שנציע ילווה, לבד מרצף המילים, גם בסימנים אחדים שיאפשרו מעקב אחרי רצף הדיבור. מצאי הסימנים הללו מבוסס בעיקרו על עבודתם של דו בואה ועמיתיו בתמלול מאגר האנגלית האמריקאית המדוברת של סנטה ברברה (דו בואה ועמיתיו, 1992; דו בואה ועמיתיו, 1993). מערך סימון זה פותח בעיקר לצורך עבודה על שיחות. במסגרת ההכנות לכינון מעמ"ד הוקלטו דוברים אחדים ברצפי זמן משתנים (יזרעאל, תשס"ג). טקסטים אחדים תומללו על פי עקרונות שיטתם של דו בואה ועמיתיו, שהותאמו לשימוש לתמלול העברית המדוברת. דוגמאות אחדות מן הטקסטים הללו פורסמו בכתב העת "לשונונו" (יזרעאל, תשס"ב, ב; ראה גם הערה 15 להלן). כאן לא נדון בסימנים אלא בגישה הכללית לתמלול הטקסט המדובר.

תחביר: היחידה המינימלית

תחביר מוגדר, בהגדרתו הפשוטה ביותר, הפופולרית, כתחום חקר הלשון העוסק בצירופי מילים (מתיוס, 1997, עמ' 368). מכאן נובע שהיחידה הקטנה ביותר של התחביר היא המילה (שם, עמ' 404). שאלה זו טעונה בירור מיוחד. מושג המילה כפי שראוי לראותו

6. למרות הכרזתם על השוני המהותי בין מבנה הלשון הדבורה לבין מבנה רבדיה האחרים של הלשון אין הצלחתם של בייבר ועמיתיו (1999) מלאה בניסיונם להשתחרר מכבלי המסורת בתיאור הלשון הדבורה (סינקליר, 2001, עמ' 354-356).

מנקודת מבט בלשנית אינו מוגדר דיו אפילו עבור הלשון הכתובה (קריסטל, 1997, עמ' 91), ובחקר הלשון המדוברת אנחנו תופשים את מושג המילה בדרך כלל כפי שמסורת הכתיבה תופשת אותו. בחקר לשונות ללא מסורת כתיבה פותחו גישות שונות לקביעת החלוקה למילים (ומי לידנו יתקע ש"מילה" הוא מושג מועיל לדיון בכל לשון באשר היא?). קריטריונים אלה, ועמם אחרים שנצפו במחקר הבלשני הכללי, עשויים להיות פונטיים או מבניים, והם כוללים בין היתר קיום סדק פוטנציאלי לאפשרות של פְּסָק בדיבור, אי אפשרות לחלוקה של רצף, אפשרות לקיום עצמאי, גבולות להתפשטות יסודות פונטיים, שיקולים של הטעמה, חילופים פרדיגמטיים ושיקולים סמנטיים (ראה למשל מתיוס, 1974, פרק 2; מתיוס, 1981, עמ' 50–59; אקמאגיאן, דמרס והרניש, 1984, סעיף 3.1; מתיוס, 1997, עמ' 404–405; קריסטל, 1997, עמ' 91; בייבר ועמיתיו, 1999, עמ' 50–54; ובמיוחד דיקסון ואייכנוולד, 2002).

ואם בתרבות שיש לה גם ערוץ כתוב עסקינן, הרי בלשון המדוברת על גווניה השונים חלוקת המילים תהיה שונה מבלשון הכתובה, ועשויה להתבסס על קריטריונים אחרים. מכל מקום, עלינו לבדוק שאלות אלה מתוך התבוננות ברצף הקולי ולא לצאת מן המקובל בחקר הלשון הכתובה. גם מסד תאורטי ונקודת המבט יכתיבו את הגדרת המילה ואת החלוקה למילים.

לאמתו של דבר, הפרדה בין מילים אינה שכיחה ברצף הדיבור. כל המאזין לדיבור בלשון זרה לו נשמע לו הדיבור כרצף בלתי פוסק, וודאי שאינו יכול להבחין בין מילים ברצף הנשמע. אנו נשוב עוד אל אותם מקומות המאפשרים בכל זאת חלוקה ליחידות, אולם אם במילים עסקינן, החלוקה אינה טריוויאלית כלל וכלל (הערות נוספות לשאלה זו ימצא הקורא המעוניין אצל יזרעאל, תשס"ג–תשס"ד). היות שמטרת מאגר הלשון לשמש תשתית מחקרית לסוגי מחקר שונים, מן הראוי להימנע ככל האפשר מהנחות מוקדמות על מבנה הלשון. לשם כך ראוי לתעתק את הרצף הדיבורי בצורה קרובה ככל האפשר לנשמע ולהגדיר את גבולות היחידות התחבריות על פי הנשמע ולא על פי שיקולים תאורטיים. אולם הדבר אינו אפשרי: כל תעתיק, ויהא הצר שבצרים, מכתוב ויתור על יסודות מסוימים של הרצף הקולי. השאלה על מה לוותר היא כמובן שאלה המערבת שיקולים תאורטיים. כפי שהערתי כבר לעיל, שיקולים מעשיים ותאורטיים הכתיבו ויתור על תעתיק פונטי למכלול הטקסטים של מעמ"ד. יש להוסיף ולומר כי קלות הקריאה אף היא שיקול חשוב, וחלוקה למילים עדיפה לאין ערוך על רישום רצף – גם בתעתיק. אנו פטורים כאן מעיון בשאלה זו, שהרי התמליל מבוסס על הכתיב המקובל ולכן גם על החלוקה המקובלת למילים. כל מחקר בשאלת מהות ה"מילה" בלשון המדוברת שירצה להתבסס על החומר שבמעמ"ד יצטרך להיזון מן ההקלטות בלבד ולהשתמש בתמליל לצרכים טכניים בלבד (כגון חיפוש יחידות והקשריהן).

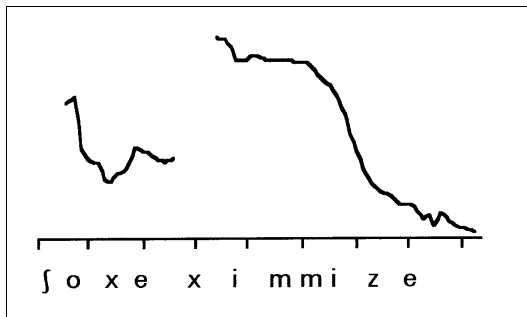
נעבור אפוא להערות אחדות לשאלת צירופי היחידות הללו ליחידות גדולות יותר ולשאלת רישומן. לא נדון כאן אלא ברמה אחת נוספת, היא רמת הקבוצה הפרוזודית (prosodic group). קבוצות פרוזודיות מתחברות אף הן להרכבים גדולים יותר, שהוצע

לכנותם "מבעים" (utterance); ראה קרסטי ומונלייה, 2005. לגישה אחרת ראה הלידיי, 1989, עמ' 82-86; הלידיי, 1994, פרקים 7, 8).

קבוצות פרוזודיות

כותרת המאמר הזה מלמדת על הצורך לצאת אל התמליל מן התאוריה. התאוריה היא המכתיבה את צורת התמליל. אפשר שיתומלל טקסט על פי המקובל ברישום השפה הכתובה, ושייתוספו לרצף המדובר סימני פיסוק. שיטה זו – הגם שלכאורה אין היא נשענת על מסד תאורטי מוצהר – מניחה במשתמע כי השפה המדוברת והשפה הכתובה דומות במבניהן. אפשר שיתומלל טקסט בשורות רציפות ובלי מרווחים או סימנים מפסיקים. כך דרכה של אילאיל יציב במחקרה על תחביר העברית המדוברת (יציב-מליבר, 2002; יציב, תשס"ב). דרך תמלול זו ננקטת כדי למנוע הנחות מוקדמות באשר למבנה השפה המדוברת (יציב-מליבר, 2000, עמ' 50-52; יציב, תשס"ב, עמ' 422-423; ראה גם בלאנש-בנווניסט וז'נז'ן, 1987; בלאנש-בנווניסט, 2000, עמ' 28-34). עם זאת שיטה זו מתעלמת במודע מיסוד מהותי במבנה הלשון המדוברת: פרוזודיה (ראה יציב-מליבר, 2002, עמ' 50; השווה בלאנש-בנווניסט, 2000, שם).

הלשון הכתובה מאורגנת בפסוקיות ובמשפטים. הלשון המדוברת מאורגנת בקבוצות פרוזודיות. הטענה הזאת, אשר נחקרה לעומקה עוד בשלהי המאה השמונה-עשרה (סטיל, 1775), פותחה ושוכללה בעשרות השנים האחרונות על ידי חוקרים רבים, בעיקר אנגלים ואמריקאים. אף כי בבסיסה דומה תפישת המושג אצל החוקרים מן האסכולות השונות, המינוח הנקוט אינו אחיד תמיד,⁷ והיבטים שונים של המושג באים לידי ביטוי



איור 1: מתאר הנגנה פשוט

בתיאורו (הלידיי, 1989, פרק 4; הלידיי, 1994, פרק 8; קרטננדן, 1997, סעיף 3.2 ופרק 4; ברזיל, 1997, פרק 1; צ'ייף, 1987; צ'ייף, 1993; צ'ייף, 1994, פרק 5; דו בואה ועמיתיו, 1992, סעיפים 4.1, 16.1-16.3 ופרק 22). גישת התמלול של מעמ"ד נשענת על התאוריה הקוגניטיבית של

7. המונחים המשמשים במחקר הם, בין היתר, intonation group, prosodic unit, tone group, intonation unit ולאחרונה בעיקר intonation unit (בעברית "יחידת הנגנה"). אני מעדיף את המונח "קבוצה פרוזודית": כפי שנראה להלן, תפישת היחידה הזאת נשענת על יסודות פרוזודיים נוספים על הנגנה.

צ'יף (1994), הרואה בקבוצה הפרוזודית את היחידה הבסיסית של הלשון המדוברת. הגישה התאורטית הזאת רואה את הקבוצה הפרוזודית כרצף דיבור שבו מתאר הנגנה יחיד ומלוכד (coherent intonation contour; צ'יף, 1994, עמ' 57–60; דו בואה ועמיתיו, 1992, עמ' 17). איור 1 מדגים מתאר הנגנה קלאסי לטקסט "שוכחים מזה"⁸. לפי האסכולה הבריטית, מתאר הנגנה יחיד מכיל, ברגיל, הבלטה (prominence) ראשית אחת בלבד (קרטנדן, 1997, סעיפים 3.4, 4.3). לא כך גישתו של צ'יף, המתירה יותר מהבלטה ראשית אחת לקבוצה פרוזודית (צ'יף, 1994, עמ' 58, המפנה גם אל לאד, 1986, שאף הוא אינו תומך בדעה השלטת; השווה גם קרטנדן, 1997, עמ' 44). אם נתבונן בניסיונות השונים להגדרת הקבוצה הפרוזודית, נבחיך כי הם מושגים בעיקר על היותה חופפת עם מתאר הנגנה יחיד, כנאמר. דא עקא, שבבואנו ללמוד מהו מתאר הנגנה, לזאת לא נמצא הגדרה של ממש, ואף התיאורים המובאים בספרות למתארי הנגנה הם, כאמור, חמקמקים. לפיכך ההתבוננות במתאר ההנגנה המלוכד, ועמו בקבוצה הפרוזודית, נעשית בעיקר על פי גבולותיו (קרטנדן, 1997, סעיף 3.2; צ'יף, 1994, עמ' 57–60; השווה גם הרינגטון וקסידי, 1999, עמ' 123–124).

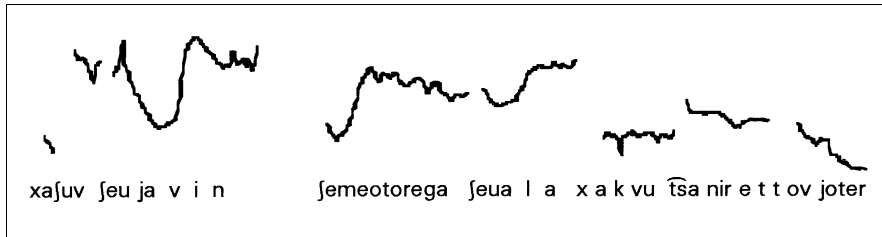
גבולות הקבוצה הפרוזודית

בקביעת גבולות הקבוצה הפרוזודית הושם עיקר הדגש על מספר קטן של קריטריונים. הקריטריונים שאנו עבדנו על פיהם עד כה מבוססים על גישת העבודה שגובשה על פי עבודותיהם של קרטנדן (1997, סעיף 3.2), של צ'יף (1993, עמ' 57–60; 1994) ושל דו בואה ועמיתיו (1992, במיוחד פרק 22; 1993). לאופר (תשמ"ז, עמ' 62–63; תשנ"ו, עמ' 289–290) הציע לראשונה את הקריטריונים הללו לעברית. גישה זו אומצה גם על ידי קבוצות המחקר של רות ברמן מאוניברסיטת תל-אביב (כהנא אמיתי וכצנברגר, 1999, עמ' 202). אלו הם:

1. קצב מזורו בראשית היחידה וקצב מואט בסופה, בעיקר הארכת ההברה האחרונה.
2. שינוי לא מדורג של גובה הקול (pitch) בין סוף היחידה לבין נקודת המוצא של יחידה חדשה.
3. פסק (pause)⁹.

8. משום שהרצף הקולי נקטע ברגיל על ידי מקטעים בלתי קוליים, גם מתאר ההנגנה איננו רציף, ומשום כך הגה בלתי קולי הפותח יחידה יימצא ברישום משמאל לתחילת קו המתאר, כבדוגמה שלפנינו. בדוגמה זו גם העיצור x השני קוטע את הרצף הקולי, בעוד שה-x הראשון מבוצע כקולי, כנראה בשל מיקומו בין תנועות ובשל קוצרו. התעתיק המשמש בדוגמאות הוא תעתיק רחב, ואינו משקף אלא במידה מועטה ביותר את הביצוע הממשי של הרצף ההגאי.
9. לצד הקריטריונים הללו קריטריון קובע בתיאור לשונות אחדות הוא דווקא שינוי מאפיין בגובה הקול בסוף היחידה (ראה הירסט ודי קריסטו, 1998, עמ' 35–36, ובתיאורי הלשונות השונות שם).

בדוגמה שלפנינו שלוש קבוצות פרוזודיות מובחנות היטב:



איור 2: שלוש קבוצות פרוזודיות בעלות מאפיינים ברורים לקביעת גבולותיהן

כיסוד איקוני בתמליל, כל קבוצה פרוזודית נרשמת בשורה נפרדת, כך שסוף שורה מסמן גבול של קבוצה פרוזודית:

חשוב שהוא יבין
שמאותו רגע שהוא הלך
הקבוצה נראית טוב יותר

גבולותיהן של הקבוצות הפרוזודיות שבדוגמה זו מובחנים בעזרת המאפיינים שזכרו:

1. קצב מזורז בראשית היחידה והארכת ההברה האחרונה: משך ארבע ההברות הראשונות ביחידה הראשונה הוא 47 מאיות השנייה (להלן: מ"ש), בעוד שמשך ההברה האחרונה לבדה הוא 38 מ"ש. משך שש ההברות הראשונות של היחידה השנייה הוא 66 מ"ש, ומשך ההברה האחרונה בה הוא 32 מ"ש. משך חמש ההברות הראשונות של היחידה השלישית הוא 74 מ"ש, ומשך הברתה האחרונה הוא 22 מ"ש. כאן יש לשים לב שההברה המובלטת ביחידה זו היא טוב, הברה הקרובה לסוף היחידה, ומשכה 28 מ"ש. מן הראוי להדגיש שעל אף שבדוגמה זו יש פער של ממש בין אורך ההברות בתחילת כל יחידה לאורך ההברות בסופה, הקריטריון המשמעותי לקביעת הגבול הוא הפער שבין אורך ההברה האחרונה של יחידה אחת לקצב הפקת ההברות הראשונות שביחידה שאחריה.

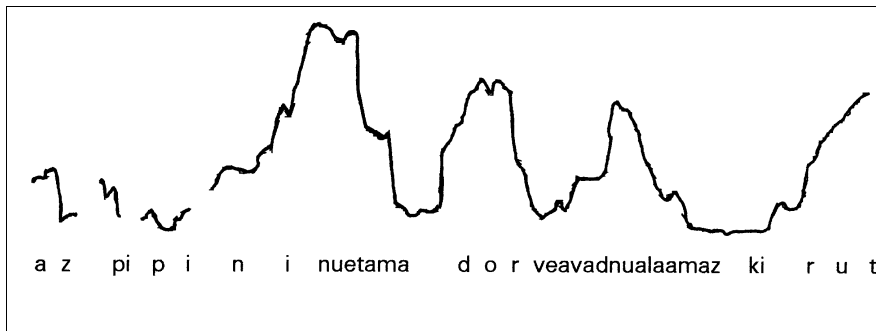
2. היערכות מחדש של גובה הקול בראשית כל יחידה: היחידה הראשונה מסתיימת בתדירות קול של 241 Hz, והיחידה השנייה מתחילה בתדירות 147 Hz – דהיינו הפרש בגובה הקול בן 94 Hz. היחידה השנייה מסתיימת בתדירות 220 Hz, והיחידה השלישית מתחילה בתדירות 137 Hz – הפרש בן 83 Hz.

3. פסק: משך הפסק בין היחידה הראשונה לשנייה הוא 18 מ"ש. בין היחידה השנייה לשלישית הפסק מינימלי (2 מ"ש) ואינו משמעותי.

לאופר מצא כי היערכות מחדש של גובה הקול היא הקריטריון החשוב בשלושה, שכן על פי ממצאיו היערכות גובה הקול לקראת נקודת מוצא חדשה תימצא בכל קבוצה פרוזודית (בלשונו: "קבוצת הנגנה" – לאופר, תשמ"ז, עמ' 63; לאופר, תשנ"ו, עמ' 289–290).

לי נראה כי לעברית הדבורה¹⁰ הקריטריון שהצבתיו לעיל ראשון, דהיינו שינוי קצב, הוא הקריטריון בעל המעמד העליון בהייררכיה לקביעת גבולותיה של קבוצה פרוזודית. מן הראוי לשים לב עוד כי מרכיב הקצב אינו מתבונן ביחידה מבחוץ דווקא, כלומר על פי גבולותיה, אלא – כפי שראינו לעיל – גם מבפנים, אל תוך היחידה, ואם אכן יאושש ממצא זה בניחות של טקסטים מגוונים ובמספר גדול הרבה יותר ניתן יהיה להשתמש ברכיב זה כדי לבנות הגדרה מהותית יותר של הקבוצה הפרוזודית (לצד קריטריונים אחרים, כגון הצטמצמות מרווח תדירות הקול [declination] או קול סדוק [creaky voice], שאינם נידונים כאן; השווה ברזיל, 1997, פרק 1; קרטנדן, 1997, סעיף 3.5; צ'ייף, 1994, עמ' 60).

כאמור, לאופר רואה בקריטריון ההיערכות מחדש של גובה הקול את הקריטריון החשוב ביותר. לאופר מוסיף ואומר כי ניסיונו מלמד שקריטריון הקצב חשוב במידה רבה, ושולי יותר קריטריון הפסק. ראינו זאת בדוגמה הקודמת, שבה הפסק בין היחידה השנייה והשלישית הוא קצר כדי כך שבלא גיבויים של הקריטריונים האחרים לא יוכר כלל באוזני בני השיח¹¹ (כדאי להעיר בהקשר זה כי פסק יכול להופיע לא רק בין קבוצות פרוזודיות אלא גם בתוך קבוצה פרוזודית; קרטנדן, 1997, סעיף 3.2.1). בדוגמה דלהלן אין כלל פסק המפריד בין שתי הקבוצות הפרוזודיות שהיא כוללת, אולם הן מובחנות יפה בעזרת שאר הקריטריונים שהוצעו, דהיינו שינוי קצב בין ההברה dor לבין ההברות שאחריה והיערכות מחדש של גובה הקול בנקודה זו.



איור 3: שתי קבוצות פרוזודיות בלא פסק ביניהן

10. הנתונים של לאופר לא היו נתונים מן העברית הדבורה (לאופר, תשמ"ז, עמ' 10; לאופר, תשנ"ו, למשל הערות 8, 10).
11. תפישת השומע פסק ממשי בדיבור תלויה באורך הפסק. יש שמפגש של סמנים אחרים נתפש כפסק, גם אם אינו קיים בפועל. לפעמים במקום פסק באה הברת מילוי או הארכת הברה מן המקטע המילולי (פייק, 1971, עמ' 68-69; קרטנדן, 1997, סעיף 3.2.1).

כדאי להעיר עוד כי הקבוצה הפרוזודית הבאה בשיח הזה אף היא אינה מופרדת מקודמתה על ידי פסק:

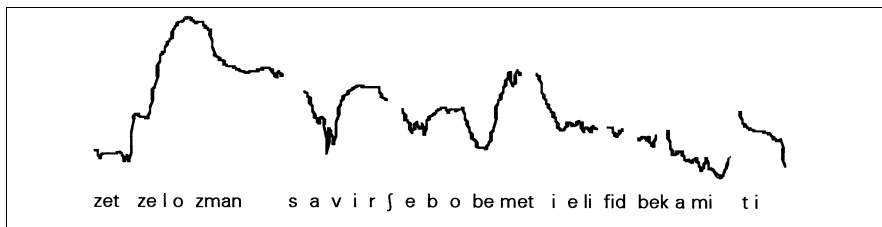
אז פי-פינו את המדור
ועבדנו על המזכירות
והבאנו מישהו מההדברה של הצבא של הקריה

הדבר אף ניכר בתוספת קוליות ל-/t/, הפונמה האחרונה בקבוצה השנייה, הבאה בין תנועה לבין ההגה הקולי [v] של וי"ו החיבור הפותחת את הקבוצה הבאה (ובתעתיק צר מן הראוי לסמן זאת).

קביעת גבולותיה של קבוצה פרוזודית היא עניין מסובך הדורש מיומנות רבה. מהירות הדיבור ואופייה המיוחד של הלשון הדבורה, לעומת הפקתו של טקסט מוכן מראש או הקראת טקסט מן הכתב, הם הגורמים לקושי העצום שבכדיקת גבולות היחידות המרכיבות את הלשון הזאת. קרטננדן (1997, עמ' 29) מבהיר:

רוב הבלשנים מניחים כי התאימות הפונטית לגבולות של קבוצות הנגנה גדולה הרבה יותר ממה שהיא במציאות. בקריאה בקול, או בהשמעת טקסט שהוכן מראש, גבולותיהן של רוב קבוצות ההנגנה אמנם מסומנים בבירור. אולם גם אצל הקוראים או הדוברים המנוסים ביותר נמצא פעמים רבות שקשה להכריע אם קיים גבול או לא. ואצל קוראים או דוברים שאינם מנוסים [...] הקשיים כפולים ומכופלים. בבואנו לבדוק דיבור ספונטני (במיוחד שיחות), חלוקה ברורה וחדה לקבוצות הנגנה אינה כה שקופה, מפאת האופי הקטוע של הדיבור הספונטני [...] ¹²

הדוגמה דלהלן תשמש לנו להמחשת הקשיים העומדים בפני המתמלל בבואו לקבוע גבולות של קבוצות פרוזודיות, וגם להמחשת המאפיינים הנתונים לנו כדי לקבעם.



אזור 4: קשיים בקביעת גבולות – עקומת תדר

12. [T]he majority of linguists assume that the phonetic correlates of boundaries between intonation-groups are far more straightforward than they actually are. In reading, or in speaking prepared texts, most intonation-group boundaries are clearly marked. But even with the most experienced readers and speakers, there are many cases where it remains difficult to decide whether a boundary is present or not. And with

תמליל רציף של הטקסט הנשמע, בלא חלוקה ליחידות, הוא זה:

זאת אומרת זה לא זמן סביר שבו באמת יהיה לי פידבק אמיתי

הקושי הראשון בפענוח הרצף הקולי היה לקבוע מה נאמר לפני הרצף זה לא זמן סביר. לאוזן בלתי מיומנת – וכך נשמע גם למתמללים בראשית עבודתם – נשמעת שם הברה הניתנת להתפרש כמילה אז, פירוש ראוי על פי ההקשר. אולם בהאזנות חוזרות פוענח הצירוף כביטוי זאת אומרת בהגייה חטופה (בת 17 מ"ש), אשר תתועתק zět.

משנקבע הדבר עמדה בפנינו גם השאלה אם הרצף הזה מהווה קבוצה פרוזודית נפרדת, דהיינו אם אפשר לקבוע גבול בינה לבין הרצף הפותח בהברה ze. התשובה שאימצו המתמללים בסופו של דבר הייתה חיובית. אולם באמת השאלה לא נפתרה. מכיוון שזו הברה יחידה, הקריטריון של האטת הקצב בסוף הקבוצה הפרוזודית מנוטרל. אף אין פסק בין ביצוע הרצף הזה לבין ההברה ze שאחריה. הקריטריון היחיד העשוי לקבוע הוא אפוא קריטריון ההיערכות מחדש של גובה הקול. האיור דלעיל אמנם מראה שינוי מדרגה בגובה הקול (38Hz). לאופר (תשמ"ז, עמ' 63) מציע שקריטריון זה קביל לשינוי בשני הכיוונים, כלומר היערכות בתדר גבוה יותר כמוה כהיערכות בתדר נמוך יותר. עם זאת לאופר לא פירט אם קיימים תנאים פרוזודיים או אחרים להעדפת אחת משתי האופציות הללו. במקרה שלנו השינוי בגובה הקול הוא מנמוך לגבוה, ומשום שהטון ממשיך ועולה אינני בטוח שמן ההכרח לפרשו כמסמן גבול בין יחידות. השאלה נותרת אפוא פתוחה.

השאלה אם אכן יש פה גבול בין קבוצות פרוזודיות או לא אינה טריוויאלית, משום שיש לה השלכות על תפישת מהות היחידות התחביריות של הלשון המדוברת. היחידה הזאת נחשבת במחקר כיחידה מסדירה (regulatory), דהיינו כיחידה שאינה מהותית למסר התוכני אלא משמשת בהסדרת השיח (צ'ייף, 1994, עמ' 63–64). בבדיקה ראשונית של כל הופעות הביטוי הזה בשיחות המוקלטות שלנו נתפש רובן המכריע כיחידה נפרדת (השווה לאנגלית האמריקאית: דו בואה ועמיתיו, 1992, סעיף 22), אולם מחקר נוסף – שיכלול את כל ההיבטים האפשריים, את התחביר, את ההקשר בשיח, את היסודות הפרוזודיים ואת הביצוע – בוודאי יעמיק עוד את ההתבוננות.

בדיקה מעמיקה של הרצף שאחרי הביטוי המסדיר הזה תלמדנו עוד על ההשלכות שיש לתמליל על המחקר ותוסיף עוד דבר מה לשאלת התאוריה המובילה לחקר התחביר. ננסה לבדוק כמה קבוצות פרוזודיות כולל הרצף הזה ולתהות על הקריטריונים

inexperienced readers and speakers [...] the difficulties are multiplied. When we consider spontaneous speech (particularly conversation) any clear and obvious division into intonation-groups is not so apparent because of the broken nature of much spontaneous speech [...]

שעל פיהם נקבע את גבולותיהן. ההכרעה היכן נקבע את הגבול בין הקבוצות נושאת עמה כמובן משמעות להבנת המבנה התחבירי ויחסו לפרוזודיה. האפשרויות הן כדלקמן:

- א. זה לא זמן סביר שבו באמת יהיה לי פידבק אמיתי
- ב. זה לא זמן סביר שבו באמת יהיה לי פידבק אמיתי
- ג. זה לא זמן סביר שבו באמת יהיה לי פידבק אמיתי
- ד. זה לא זמן סביר שבו באמת יהיה לי פידבק אמיתי

בשמיעה נדמה כי כל אחת מארבע המילים זמן, סביר, שבו ובאמת נהגית בגובה קול שונה מן המילה הקודמת. אולם השינוי בין לא לבין זמן נתפש כהדרגתי, דהיינו כשינוי פנימי של מתאר ההנגנה, בירידה מן המילה המובלטת לא. כך הוא בוודאי גם ברצף שלאחר המילה באמת, הנתפשת כמילה המובלטת בקבוצה פרוזודית שנייה ברצף זה (כך על פי אפשרויות ב, ג וד), הבלטה שלאחריה יש ירידה בגובה הקול הנמשכת כמעט עד סוף הקבוצה. יחידה זו מסתיימת בטון ממשך, בעליית תדירות הקול בהברה האחרונה, עלייה המסמנת כי הקבוצה הזאת אינה מסיימת את המבע (ראה דו בואה ועמיתיו, 1992, סעיף 6.2; דו בואה ועמיתיו, 1993, עמ' 52–54; קרטנדן, 1997, סעיף 4.4.3. לעברית ראה חן, תשל"ב, עמ' 299–300; לאופר, תשמ"ז, סעיפים 2.7, 2.8, 2.10. [=תבניות הנגנה 6, 7, 9]). גם עליית גובה הקול בהברה האחרונה של המילים סביר ושבו אפשר שתיתפש כמסמנת הטעמת המשך.

לשלושת הרצפים סביר, שבו ובאמת מבנה הנגנה דומה: תדירות קול נמוכה ואחריה תדירות קול גבוהה. לכאורה אפשר להניח שתדירות הקול הנמוכה שבהברה הראשונה של כל מילה מן המילים הללו יכולה להיתפש כפתיחה של קבוצה פרוזודית, שהרי היא באה לאחר תדירות קול גבוהה, במבנים שכמותם כבר ראינו לעיל. ירידות ועלויות הן אמנם מהות ההנגנה, אולם גם מסגרת המילה בנויה מהברות מוטעמות ובלתי מוטעמות, מה שנתפש כקצב של הדיבור (rhythm). קצב הדיבור נשלט אפוא על ידי גורמים דומים לאלה היוצרים את ההבלטות: תדר, עצמה ומשך (קרטנדן, 1997, סעיף 1.3 ופרק 2; הירסט ודי קריסטו, 1998, עמ' 3–13; לעברית ראה אנוך וקפלן, תשכ"ט; למשך התנועות המוטעמות בעברית ראה טנא, תשכ"ב; השווה קוריאגוזה, 1994). דרוש עוד מחקר רב לעניין יחסי הגומלין בין קצב לבין הנגנה בעברית. ירידת התדר בתחילת המילה שבו (286 Hz^{-1} ל- 235 Hz), כמשתקף בבירור מן הגרף לדליל, איננה כה גדולה כירידת התדר בזוג המילים שלפניה, זמן וסביר (מ- 308 Hz ל- 216 Hz), אולם היא דומה לזו שבין שבו לבין המילה שאחריה, באמת (מ- 260 Hz ל- 219 Hz). תדירות הקול במילים סביר ובאמת מגיעה בערך אל תדר הפתיחה של היחידה זאת אומרת או של זה. יש לשים לב לצמצום הדרגתי של מרווח התדרים לאורך המקטע כולו, צמצום האופייני

הן לקבוצה פרוזודית יחידה והן לרצפים של קבוצות פרוזודיות המאוגדות יחדיו (קרטננדן, 1997, סעיפים 3.5, 4.4.4.4, 5.5.1; לאופר, תשמ"ז, עמ' 64; השווה הירסט ודי קריסטו, 1998, עמ' 1-20).

נבדוק גם את שני הקריטריונים האחרים: משך ההברות והפסקים. פסקים אין כאן כלל. משך ההברה vir 28 מ"ש, משך ההברה bo 22 מ"ש, ומשך ההברה ti של סוף המקטע כולו הוא 23 מ"ש. דא עקא, שהשוואה למשך הביצוע של ההברות הסמוכות אינה מובהקת, לבד מקצב מהיר בראשית היחידה הראשונה ובמילה יהיה. ובכן, אין סמנים מובהקים גם בקריטריון הקצב. מכיוון שעליי להגיש תמליל יחיד, עליי להכריע בין האפשרויות השונות. אם נשוב לתחילת הרצף הזה ונוסיף את השאלה אם הביטוי זאת אומרת מהווה קבוצה פרוזודית נפרדת או לא, מספר האפשרויות גדל אף יותר. ההכרעה שלי היא:

זאת אומרת

זה לא זמן סביר

שבו באמת יהיה לי פידבק אמיתי

והיא הכרעה שאני כופה על הקורא באי רצון.¹³

מחקר מעמיק, שיבדוק תפישה שמיעתית בשילוב ניתוח אקוסטי הכולל את כל מרכיבי הקול: משך, תדירות ועצמה (שאלה שלא נגענו בה כאן), ובשילוב שיקולים של סוג ההברה ומהות ההגאים היחידים (השווה מיקסדורף, 1998, סעיף 4.2.2) עשוי ללמדנו על המבנה הפרוזודי של העברית ולסלול דרכים חדשות למחקר התחבירי של העברית המדוברת בכלל ושל העברית הדבורה בפרט.¹⁴ שהרי תפקידנו הוא למצוא קריטריונים פורמליים שבעזרתם נוכל לנתח את הלשון המדוברת, כדי להבין את מבניה. בהנחה שהקבוצות הפרוזודיות הן היחידות הבסיסיות של הלשון המדוברת, תפקידנו להגדיר את היחידות האלה ולבסס קריטריונים פרוזודיים לתיאור המבנים הללו ולקביעת גבולותיהם. הדיון דלעיל וממצאיו מדגימים את חשיבות העניין. ממצאי הניתוח מתבססים על שיקולים תאורטיים ומובילים לניתוח מבנה הלשון ומכאן לתאוריות נוספות בתחומי עיון אחרים. במקרה שלנו תעמוד השאלה של היחס התחבירי בין הצירוף שבו לרצף שאחריו

13. בתמליל הטקסט הזה כפי שפורסם (יזרעאל, תשס"ב, ב, עמ' 293) חלוקה שונה:

זאת אומרת

זה לא זמן סביר שבו

באמת יהיה לי פידבק אמיתי

תמליל זה נערך בלי ניתוח אקוסטי והתבסס רק על האזנה.

14. צעד ראשון בכיוון זה נעשה על ידי ורד זילבר-ורוד (תשס"ה). ראה גם יזרעאל וזילבר-ורוד, ברפוס. להערות על היחס בין תפישה לרישום אקוסטי ראה לאופר, תשמ"ז, עמ' 32-33, הערות 15-18.

ובין הרצף הפותח בצירוף שבו לבין הרצף שלפניו ושל המעמד התחבירי-סמנטי של המרכיבים השונים שבהרכב.¹⁵

ודאי, דרוש עוד מחקר מקיף בשאלות אלה, גם בשאלת הפרוזודיה וגם בשאלות אחרות. למשל השאלה מהם יחסי הגומלין בין קבוצות פרוזודיות לבין התחביר והסמנטיקה היא שאלה שיש לתת לה עדיפות בחקר הלשון המדוברת: האם לתחביר המסורתי יש מקום בתיאור השפה המדוברת? היש טעם לדבר על משפט? על פסוקית? מחקר לשונות אחרות שולל חד-משמעית את תקפותו של המושג "משפט" בלשון המדוברת, אולם הדעות חלוקות לגבי מעמדה של הפסוקית (ראה למשל הלידי, 1994, פרק 8; צ'יף, 1994, עמ' 56-59; מילר וויינרט, 1998, פרק 2; בלאנש-בונוויסט ועמיתה, 1990, עמ' 173-175).

דיבור, תמליל ותחביר

אחרי שטעמנו מן הנושא על קצה המזלג וראינו מה לתאוריה ולתמליל, נתבונן עתה בשאלות אחרות העומדות בפני המתמלל בבואו להעביר את הדיבור אל הערוץ הכתוב למטרת מחקר. כאן מן הראוי לפתוח בחזרה על העיקרון הפרלימינרי, הפרימיטיבי, שחייב לעמוד כנר לרגלינו בבואנו להגיש תמליל (או תעתיק) של טקסט למשתמש הפוטנציאלי:

הטקסט הכתוב – במגבלות הפישוט וההפשטה שנהגנו בדיוננו כאן – צריך לשקף את מה שנאמר בפועל. הדבר נשמע מובן מאליו, אולם אין זה כך. הדגמתי לעיל אך מעט מן הקשיים התאורטיים העומדים בפני המתמלל. אולם מלבד הללו ודומיהם, המערכת הלשונית-חברתית מציבה בפני הרושם עוד שני מכשולים ממשיים:

1. השאיפה ללשון סטנדרטית.
2. המשיכה ללשון הפרטית, או אם תרצו: שתלטנותה של הלשון הפרטית.

15. ההרכב הפרוזודי אינו מסתיים במילה אמיתי, כפי שמלמד הטון הממשיך. ההרכב כולו הוא זה:

זה לא זמן סביר |

שבו באמת יהיה לי פידבק אמיתי |

.. כולל |

.. ראייה נכונה |

.. על מנהל שלי ||

הסימנים המשמשים כאן: סוף שורה = גבול קבוצה פרוזודית; | = גבול קבוצה פרוזודית בטון ממשך (בדרך כלל בעלייה או בירידה קטנות); || = גבול קבוצה פרוזודית בטון מסיים (בדרך כלל בירידה גדולה); \ = גבול קבוצה פרוזודית בטון של פנייה (טון עולה; בדרך כלל לשאלה שאין בה מילת שאלה); -- = קבוצה פרוזודית קטועה; רווח = גבול בין מילים; - = מילה קטועה; .. = פסק קצר; [שלום] = חפיפה בדיבור; אה {e}, אה {a} = היסוס בדיבור; (שלום) = תמליל לא בטוח; @ = הברה לא מזוהה. ראה עוד יזרעאל, תשס"ב, ב, עמ' 290-291.

לשון נורמטיבית ולשון סטנדרטית

בדברי על לשון סטנדרטית אין כוונתי ללשון התקנית, הנורמטיבית, דווקא אלא לסטנדרט המקובל. אולם גם הלשון הנורמטיבית – שאינה סטנדרטית – עשויה להטעות את רושם התמליל. למשל, המתמלל עשוי להעביר אל הכתב צורות כינויים או צורות פועל ברבים המסתיימות ב־m ברישום בנו"ן סופית דווקא. הטעות הזו קורית גם מתמללים מנוסים, שכן תיקונים כאלה אינם נעשים בהכרח במודע. להלן תמליל (ראשוני, לא מעובד) שנעשה על ידי מתמללת בעלת מודעות:

מה
אתן לא חושבות
מה אתן דפוקות
התעצבן

צורת הכינוי נהגתה בשתי ההיקריות ב־m (atem), וזו צורה צפויה בהחלט, מן הסתם חלק מן הסטנדרט השולט בדיבור העברי. עם זאת התקן, אשר המתמללת נוקטת בוודאי בכתיבתה, הסיט אותה מן המקובל בדיבור ומן הנשמע בפועל אל הסטנדרט הכתוב.

הדוגמה דלעיל מציגה כשל העשוי להשפיע על הניתוח המורפולוגי. הדוגמה דלהלן מציגה כשל העשוי להשפיע על הניתוח התחבירי. דוגמה זו מציגה תיקון הנובע מהעברה לסטנדרט שאינו מוכתב בהכרח:

אז האישה
שהיא מחכה שתבוא לגור אתה
האשת חיל
האישה שהיא צריכה לבוא

ובאמת יש פה כינוי נסתרת נוסף:

אז האישה
שהיא מחכה שהיא תבוא לגור אתה
האשת חיל
האישה שהיא צריכה לבוא

התמליל הראשון מושפע מסטנדרט אחר, שונה מן הסטנדרט הלשוני של הדוברת המסוימת הזאת, בוודאי קרוב יותר אל הסטנדרט הלשוני של העברית המדוברת בפי המתמללת.

מכשלות אחרות

גם אם אנו מודעים לצורך להעביר במדויק את הנשמע ולא לתקנו, וגם אם נימלט מן הספונטניות שבתמלול, יש מוקשים רבים בדרך מהדיבור הנשמע אל הטקסט הכתוב.

הדרך שהטקסט המדובר עובר מרגע יצירתו ועד לרישומו ולהבנתו מתוך הטקסט החזותי ארוכה ורצופת מעברים. בצורה פשטנית למדי אפשר לתאר אותה כך:

מסר < נסיבות הדיבור < דיבור < תוֹךְ סביבתי (רעש) < תיווך אלקטרוני < שמיעה < פענוח < רישום < קריאה < מסר

לא כל הנאמר נתפש כמות שהוא. ידועה האמרה כי עמימות, או דר-משמעות, קיימת אך ורק אצל השומע, לא אצל הדובר. הדובר יודע מה רצה לומר, לא כן השומע. עמימות נוצרת מסיבות רבות, לשוניות וחוג-לשוניות. עמימות מובנית בהומונימיה, בהקשר לשוני, בהקשר תרבותי, בהקשר אישי ספציפי, בנתק תרבותי או חברתי, בנתק פיזי, בערוץ התקשורת, בסביבה רועשת ובתרבות הדיבור. לענייננו, מעבר הדרך האחרון הוא ממשק רושמי התמליל והמשתמשים בו. אביא שתי דוגמאות כדי להמחיש רק מעט מן המכשולים הללו ומתוצאותיהם למחקר הלשון.

הדוגמה הראשונה תראה כי העברה בלתי מדויקת אל הערוץ החזותי עשויה לגרום לעיוות בתפישה התחבירית. מתמללים שונים תמללו אותו קטע בצורות שונות. הטקסט כפי שתומלל תמלול אחרון הוא כדלקמן:¹⁶

סיגל: רציתי שתיקחי פה ימינה |
לצאת דרך אה {e} --

מרב: לא משנה |
אז ניסע דרך אחרת |
אה {a} |
מה את לחוצה ||
בעלך לא גר --
הוא לא בבית |
ואת [ישנה אצלי] ||

סיגל: [לא את זה] ||
בשבילך ||

דיבורה של מרב נבלע ברעש באמצעו, ומשום כך נחלקו המתמללים – ארבעה במספר – בניסיונם לפענח את הקטע המעורפל, כדלקמן:

א: אז ניסע דרך הבית הירוק
מה את לחוצה

ב: אז ניסע דרך אה {e}
מה את לחוצה

ג: אז ניסע דרך (אחרת)
אה {a}
סיגל: מה את לחוצה

16. לביאור הסימנים ראה הערה 15. השמות בדויים כדי לשמור על אנונימיות המוקלטים.

1ד: אז ניסע דרך @
מה את לחוצה

2ד: אז ניסע דרך @ סיגלי
מה את לחוצה

3ד: אז ניסע דרך אחרת
סיגלי
מה את לחוצה

תמליל א מציג פענוח שאינו סביר (לפחות בדיעבד), אולם שנבנה על פי דגם תחבירי סטנדרטי. תמלילים ב ו-ג הושפעו בוודאי מהרצף הקודם, שבו אמרה סיגל, בת שיחה של מרב, לצאת דרך אָה. בתמליל ג מסומן גם חילוף של דוברת בניגוד לאבחנת המתמללים האחרים, שלא הבחינו בחילוף כזה. תמליל ד1 זהיר יותר, ומציג הברה יחידה שלא פוענחה. בתמליל מאוחר יותר של אותה מתמללת (2ד) נדמה לה כי שמעה – באותה קבוצה פרוזודית – גם פנייה לסיגל. היא לא ויתרה עדיין על הברה נוספת, שנותרה בלתי מפוענחת, לפני אזכור השם. בהאזנה נוספת נתפענח לה הסתום כשם התואר אחרת, הבא על פי התמליל לפני הפנייה בשם לסיגל, והפעם נשמעה הפנייה למתמללת לאחר גבול הקבוצה הפרוזודית. התמליל הסופי כפי שהוצג לעיל, שנערך על ידי מתמלל אחר, החליף את השם סיגלי בשאלה אָה (השווה תמליל ג), שהוצבה בסוגריים כדי לסמן תמליל לא בטוח.

בתמליל הסופי, כמוהו כתמלילים ג ו-ד3, תחביר היחידה ניסע דרך אחרת שונה באורח מהותי מן התחביר המוצג בשאר התמלילים: המילה דרך אינה מתפקדת בו כמילת יחס אלא כשם עצם במעמד של משלים ישיר לפועל (שאת מעמדו ראוי לבדוק):

ניסע | דרך | הבית | הירוק
ניסע | Ø | דרך | אחרת

מבנה זה אינו מבנה שמקובל לראותו תקני בלשון הכתובה. הוא מצוי הרבה בלשון הדבורה, ומן הראוי לבדוק מה מקומו במבנה הלשון המדוברת על מערך חלופותיה. גם הדוגמה דלהלן ממחישה את הצורך בהעברה מדויקת של הנשמע. כאן חפיפה קדמית (דיבור במקביל שגבר על דיבורו של הדובר המתומלל) הפריעה להאזנה.

אתה חייב כי ככה תפרט את ה
מנהלים

האזנות נוספות של הקטע הניבו את התמליל הבא:

אתה חייב
כי ככה אתה מפרט את המנהלים

תמליל זה שונה מקודמו בשני היבטים: בהיבט המילולי ובחלוקה לקבוצות פרוזודיות. דיוק בתמליל ייטיב את הניתוח לא רק מבחינת התחביר (האם המבנה הפותח בכי בא

בקבוצה פרוזודית אחת עם המשפט אתה חייב, והאם יש גבול פרוזודי בין תווית היידוע לבין המילה מנהלים), אלא גם מבחינת מערכת הזמנים (תפרט לעומת אתה מפרט).

הנה כי כן, רבות המכשלות, אולם רבים הסיכויים. מסד תאורטי מוביל לתמליל ראוי, ותמליל ראוי פותח פתח לניתוח תחבירי שיוכל להישען על בסיס איתן. ברשימות אלה העליתי שאלות ובעיות, לא פתרונות. בלא מחקר בסיסי מעמיק המושתת על מסד נתונים רחב – לא ייתכנו מסקנות ראויות על מבנה הלשון. מחקר הלשון המדוברת עדיין בראשיתו, והשאלות הבסיסיות עדיין ניצבות לפנינו. אנו מקווים כי מאגר העברית המדוברת בישראל יהווה מסד נתונים ראוי, כך שיוכל לשמש בסיס לבדיקה של שאלות מחקר ולאישוש או הפרכה של השערות מחקר.

הקיצורים הביבליוגרפיים

- K. Ehlich, "Phrase averbale, phrase nominale? La constellation = 2005", *Syntaxe & Sémantiques 6: Aux marges de la prédication*, Caen 2005, pp. 103–124
- E. Ochs, "Transcription as Theory", *Developmental Pragmatics*, = 1979, ed. E. Ochs and B. B. Schieffelin, New York 1979, pp. 43–72
- אנוך וקפלן, תשכ"ט = פ' אנוך וג' קפלן, "המהות הפיסיקלית של ההטעמה בעברית הישראלית", לשוננו לג (תשכ"ט), עמ' 208–222
- A. Akmajian, R. A. Demers, and R. M. Harnish, *Linguistics: An Introduction to Language and Communication*², Cambridge, MA 1984
- בורוכובסקי-בר־אבא, תשס"ב = א' בורוכובסקי-בר־אבא, "מן הכתב אל הדיבור: שמות סימני הפיסוק בפונקציות שונות בלשון", לשוננו סד (תשס"ב), עמ' 102–136
- D. Biber, S. Johansson, G. Leech, S. Conrad, and E. Finnegan, *Longman Grammar of Spoken and Written English*, Harlow 1999
- C. Blanche-Benveniste, *Approches de la langue parlée en français*, Gap–Paris 2000
- C. Blanche-Benveniste and C. Jeanjean, *Le français parlé: Transcription et édition*, Paris 1987
- C. Blanche-Benveniste, M. Bilger, C. Rouget, and K. van den Eynde, *Le français parlé: Études grammaticales*, Paris 1990
- D. Brazil, *The Communicative Value of Intonation in English*, = 1997 [1985] Cambridge

- J. Du Bois, "Transcription Design Principles for Spoken = 1991, דו בואה, Discourse Research", *Pragmatics* 1 (1991), pp. 71–106
- J. W. Du Bois, S. Cumming, S. Schuetze-Coburn, and = 1992, דו בואה ועמיתיו, D. Paolino, *Discourse Transcription*, Santa Barbara 1992
- J. W. Du Bois, S. Schuetze-Coburn, S. Cumming, and = 1993, דו בואה ועמיתיו, D. Paolino, "Outline of Discourse Transcription", *Talking Data: Transcription and Coding in Discourse Research*, ed. J. A. Edwards and M. D. Lampert, Hillsdale, NJ 1993, pp. 45–89
- R. M. W. Dixon and A. Y. Aikhenvald (eds.), *Word: = 2002, דיקסון ואייכנוולד, A Cross-linguistic Typology*, Cambridge 2002
- D. Hirst and A. Di Cristo (eds.), *Intonation Systems: = 1998, הירסט ודי קריסטו, A Survey of Twenty Languages*, Cambridge 1998
- M. A. K. Halliday, *Spoken and Written Language*², Oxford 1989 = 1989, הלידיי
- M. A. K. Halliday, *An Introduction to Functional Grammar*², = 1994, הלידיי, London 1994
- B. Hary and S. Izre'el, "The Preparatory Model of the = 2003, הרי ויזרעאל, Corpus of Spoken Israeli Hebrew (CoSIH)", *Corpus Linguistics and Modern Hebrew: Towards the Compilation of The Corpus of Spoken Israeli Hebrew (CoSIH)*, ed. B. H. Hary, Tel-Aviv 2003, pp. 189–219
- J. Harrington and S. Cassidy, *Techniques in Speech = 1999, הרינגטון וקסידי, Acoustics*, Dodrecht 1999
- זילבר-רודר, תשס"ה = ו' זילבר-רודר, מאפייני גבולות של יחידות פרוזודיות בעברית הדבורה: ניתוח תפסתי ואקוסטי, עבודת גמר, אוניברסיטת תל-אביב, תשס"ה (http://www.openu.ac.il/Personal_sites/download/Characteristics_of_Prosodic_Unit_Boundaries.pdf)
- חן, תשל"ב = מ' חן, "מבטאה של העברית הישראלית", לשוננו לו (תשל"ב), עמ' 219–212, 287–300
- טנא, תשכ"ב = ד' טנא, "המשך הנמדד של התנועות בעברית", לשוננו כו (תשכ"ב), עמ' 200–268
- יזרעאל, תשס"ב, א = ש' יזרעאל (עורך), בסיועה של מ' מנדלסון, מדברים עברית: לחקר הלשון המדוברת והשונות הלשונית בישראל (תעודה, יח), תל-אביב תשס"ב
- יזרעאל, תשס"ב, ב = ש' יזרעאל, "מאגר העברית המדוברת בישראל (מעמ"ד): דוגמאות טקסטיות", לשוננו סד (תשס"ב), עמ' 289–314
- יזרעאל, תשס"ג = ש' יזרעאל, "מאגר העברית המדוברת בישראל (מעמ"ד), שלב א': בדיקה טרומית – דר"ח ראשוני", קול ליעקב: אסופת מאמרים לכבוד פרופ' יעקב בן-טולילה, בעריכת ד' סיון ופ"י הלוי-קירטצ'וק, באר-שבע תשס"ג, עמ' 211–222

- יזרעאל, תשס"ג-תשס"ד = ש' יזרעאל, "מחקר העברית המדוברת: הצעד הראשון", לשוננו לעם נד (תשס"ג-תשס"ד), עמ' 106-119
- S. Izre'el, "Transcribing Spoken Israeli Hebrew: Preliminary = 2004 Notes", *Perspectives on Language and Language Development*, ed. D. Ravid and H. Bat-Zeev Shyldkrot, Dordrecht 2004, pp. 61-72
- יזרעאל, הרי ורהב, תשס"ב = ש' יזרעאל, ב' הרי וג' ורהב, "לקראת כינון מאגר העברית המדוברת בישראל", לשוננו סד (תשס"ב), עמ' 265-287
- יזרעאל וזילבר-וורוד, בדפוס = ש' יזרעאל וו' זילבר-וורוד, "אומר לנתח לנתח: על תפישת הקבוצה הפרוזודית בעברית המדוברת", בלשנות עברית 63-64 (בדפוס)
- S. Izre'el and G. Rahav, "The Corpus of Spoken = 2004 Israeli Hebrew (CoSIH), Phase 1: The Pilot Study", *LREC 2004 Sattelite Workshop: Fourth International Conference on Language Resources and Evaluation: Compiling and Processing Spoken Language Corpora* (Lisbon, Portugal), ed. N. Oostdijk et al., Paris 2004, pp. 1-7
- יציב, תשס"ב = א' יציב, "מתמלול של שפה מדוברת לניתוח בסבכה", בתוך יזרעאל, תשס"ב, א, עמ' 421-436
- I. Yatziv-Malibert, *Méthodologies pour la description de = 2002 quelques phénomènes syntaxiques de langue parlée: Application à l'hébreu moderne*, Ph.D. dissertation, École Pratique des Hautes Études, Sciences historiques et philologiques, Paris 2002
- D. Cahana-Amitay and I. E. Katsenberger, = 1999 "Intonation Units as Processing Units in Spoken Discourse", *Developing Literacy across Genres, Modalities, and Language*, ed. D. A. Aisenman, Tel-Aviv 1999, pp. 199-207
- D. R. Ladd, "Intonational Phrasing: The Case for Recursive = 1986 Prosodic Structure", *Phonology Yearbook* 3 (1986), pp. 311-340
- לאופר, תשמ"ז = א' לאופר, הנגנה, ירושלים תשמ"ז
- לאופר, תשנ"ו = א' לאופר, "הפסקה בדיבור רצוף ופיסוק", מחקרים בלשון העברית ובלשונות היהודים מוגשים לשלמה מורג, בעריכת מ' בר-אשר, ירושלים תשנ"ו, עמ' 277-294
- J. Miller and R. Weinert, *Spontaneous Spoken Language: = 1998 Syntax and Discourse*, Oxford 1998
- H. Mixdorff, *Intonation Patterns of German: Model-Based = 1998 Quantitative Analysis and Synthesis of F0 Contours*, D.Eng. thesis, TU Dresden 1998 (<http://www.tfh-berlin.de/~mixdorff/thesis/index.html>)
- P. H. Matthews, *Morphology: Introduction to Theory of Word = 1974 Structure*, Cambridge 1974

- P. H. Matthews, *Syntax*, Cambridge 1981 = 1981, מתיוס,
- P. H. Matthews, *The Concise Oxford Dictionary of Linguistics*, = 1997 מתיוס,
Oxford 1997
- J. Steele, *An Essay towards Establishing the Melody and Measure = 1775*, סטיל,
of Speech, to be Expressed and Perpetuated by Peculiar Symbols, London
1775 (reprinted as no. 172 in the Series English Linguistics 1500–1800,
Menston, England 1969)
- J. Sinclair, *International Journal of Corpus Linguistics* 6 (2001) = 2001, סינקלייר,
[ביקורת על בייבר ועמיתיו, 1999] pp. 339–359
- K. L. Pike, “General Characteristics of Intonation”, *Intonation*, = 1971, פייק,
ed. D. Bolinger, Harmondsworth 1971, pp. 53–82 (= K. L. Pike, *The
Intonation of American English*, Ann Arbor 1945, pp. 20–41)
- W. Chafe, “Cognitive Constraints on Information Flow”, = 1987, צ׳׳יף,
Coherence and Grounding in Discourse, ed. R. S. Tomlin, Amsterdam 1987,
pp. 21–51
- W. Chafe, “Prosodic and Functional Units of Language”, = 1993, צ׳׳יף,
Talking Data: Transcription and Coding in Discourse Research, ed. J. A.
Edwards and M. D. Lampert, Hillsdale, NJ 1993, pp. 33–43
- W. Chafe, *Discourse, Consciousness, and Time: The Flow and = 1994, צ׳׳יף,
Displacement of Conscious Experience in Speaking and Writing*, Chicago
1994
- F. Kuriyagawa, “The Vocalization Signs and Actual Length = 1994, קוריאגוואה,
of Vowels in Hebrew”, *Proceedings of the Eleventh World Congress of
Jewish Studies, Jerusalem, June 22–29, 1993, Division D, Volume 1: The
Hebrew Language, Jewish Languages*, Jerusalem 1994, pp. 49–56
- A. Cruttenden, *Intonation*², Cambridge 1997 = 1997, קרטננדן,
- D. Crystal, *The Cambridge Encyclopedia of Language*², = 1997, קריסטל,
Cambridge 1997
- E. Cresti and M. Moneglia (eds.), *C-ORAL-ROM: = 2005, קרסטי ומונלייה,
Integrated Reference Corpora for Spoken Romance Languages (Studies in
Corpus Linguistics, 15)*, Amsterdam 2005