

לְשׁוֹנֵנוּ

כרך שישים ושישה, חוברת שלישית ורביעית
התשס"ד

תדפיס

לאנשי (גמס"ג), "גמטיק נושי א זכיוג
מקורו"



הוצאת האקדמיה ללשון העברית

תעתיק פונטי של עברית מדוברת

תגובה למאמרו של ש' יורעאל, "מאגר העברית המדוברת בישראל (מעמ"ד):
דוגמאות לטקסטים", לשוננו סד (תשס"ב), עמ' 289-314

אחת המטרות של תעתיק פונטי היא הצגה גרפית של דיבור. אפשר לדרג תעתיקים שונים של קטע דיבור אחד בסולם שבתחתיתו התעתיק הרחב ביותר ובראשו התעתיק הצר ביותר. בתעתיק הרחב ביותר מסמנים רק את אבני הבניין הלשוניות של שפת המדבר, את הפונמות, ולמעשה התעתיק הפונטי הרחב ביותר הוא תעתיק פונמי. לעומתו, בתעתיק הפונטי, המוצב בראש הסולם, מציינים את כל הגוונים שבדיבור, גם את אלה המאפיינים את זהותו של המדבר ואת מצב רוחו. התעתיק הצר ביותר האידאלי הוא הצגה גרפית נאמנה, שוות ערך להקלטה עצמה.

חוקרים משתמשים בשלבים שונים בסולם התעתיק, לפי צורכיהם ויכולתם. לשם זיהוי של דובר משתמשים בתעתיק צר מאוד. לעומת זאת, לרוב צורכי הבלשנות אפשר להסתפק בתעתיק רחב מאוד, שהוא זהה לתעתיק פונמי שמצוינות בו רק יחידות הבניין המהותיות לשפה הנבדקת. כאמור, הצרכים השונים מכתבים את סוג התעתיק. אם מטרתנו לגלות את אוצר המילים או את שכיחות המילים בשפה, נוכל להסתפק בתעתיק רחב. אך אם נרצה, למשל, לציין הבדלי הגייה דיאלקטיים או אישיים, לא נוכל להסתפק בתעתיק פונמי ויהא עלינו להשתמש בתעתיק צר יותר.

במאמרו הציג שלמה יורעאל את מאגר העברית המדוברת בישראל (מעמ"ד) וכלל בו שני תעתיקים לדוגמה. בדבריי אלו ארצה לדון בשיטת התעתיק שהוצגה שם, ומדבריי יובן גם מה נדרש לדעתי בתעתיק פונטי של עברית מדוברת. אפתח בתשבחות על שיטת התעתיק שנבחרה ואנמק מדוע זוהי הבחירה המוצלחת, אך אמשיך ואראה כי התעתיק המודגם אינו משתמש כראוי בשיטה שנבחרה. אסיים בהצבעה על החסר בתעתיק ובתמליל.¹

1. בעניין "תכתוב ותמלול" העירה רונית גדיש בלשוננו (סה [תשס"ג], עמ' 97), ואני מודה לפרופ' משה בראשר על שהפנני להערתה. מליאת האקדמיה קבעה בתשנ"ה את שם הפעולה כתכות לציון ההעברה של טקסט מוקלט לכתב. מכיוון שבדבריי אלה אדון במאמרו של יורעאל, אשתמש במונחים שלו.

שיטת התעתיק של יפ"א

אני מבקש לברך את מתכנני המאגר על החלטתם לתעתק את המאגר בתעתיק הפונטי הבינלאומי בשיטת יפ"א (IPA – ראשי התיבות של International Phonetic Alphabet). מכיוון שלצערנו יש חוקרי לשון בארץ שאינם משתמשים בתעתיק הפונטי הזה, ננצל הזדמנות זו כדי להציג את מעלותיו.

הבחירה בסימני יפ"א היא תוצאה של תהליך ממושך שראשיתו בשנת 1886, בהקמת האגודה הפונטית הבינלאומית. סימני יפ"א נבחרו בקפידה, והקווים המנחים היו פשוטות בסימון מצד אחד ובחירה בסימנים שהם אינטואיטיביים לרוב החוקרים בעולם מצד שני. מעלה נוספת של סימני יפ"א היא מספרם המועט והעובדה שכל אחד מהם מוגדר בצורה חד-משמעית מבחינה מדעית. לוח יפ"א הוא בן עמוד אחד בלבד, אך עמוד תמים זה מסכם את הסימנים שהתקבלו על ידי קהילת הבלשנים, ומבחינה מסוימת הוא תמצית מדע הפונטיקה. כל היודע להשתמש בלוח יכול לסמן או לפענח כל הגה מבחינת הקיים בכל שפה משפות בני האדם. יתר על כן, די בעמוד האחד של לוח יפ"א כדי לתעתק (או לפענח) בכל דרגת דיוק הנדרשת למתעתק. הלוח מאפשר לסמן מחד גיסא תעתיק פונטי רחב ביותר, תעתיק פונמי, ומאידך גיסא אותו לוח עצמו מאפשר גם תעתיק פונטי צר ככל שמאפשרים אופן המתעתק או מכשיריו. בעזרת הלוח אפשר לסמן גם את רוב ההגאים בלשונם של הלוקים בדיבורם. היודעים להשתמש בלוח יפ"א יכולים לסמן ולפענח גם תופעות על-מקטעיות כמו דרגות טעם, גבולות (של הברה, של קבוצת משמעות ושל קבוצת הנגנה), משך, טונמות והנגנה. סיכומו של דבר, שיטת התעתיק של יפ"א היא שיטה מדעית בין-לאומית, המאפשרת לסמן בצורה גרפית את רוב תופעות הלשון שבכל שפות העולם, ובכלל זה את הדיבור של הלוקים בדיבורם.

לוח יפ"א המעודכן הוא תמצית של ידע פונטי שהצטבר במשך שנים רבות. סיכום מעודכן של הוראות השימוש בו פורסם בספר יפ"א.² בקובץ ההנחיות אין מצוין שם מחבר, כי הוא פרי עבודה משותפת של פונטיקאים רבים במשך דורות, כמוסבר שם בהקדמה. לוח יפ"א המעודכן האחרון יצא בשנת 1996, והוא מפורסם בכל ספר בסיסי בפונטיקה שיצא מאז. אפשר להעתיק חינם את הלוח המעודכן מאתר האינטרנט של האגודה הפונטית הבינלאומית: <http://www2.arts.gla.ac.uk/IPA/ipa.html>³.

2. IPA (=International Phonetic Association), *Handbook Of the International Phonetic Association: A Guide to the Use of the International Phonetic Alphabet*,

ed. J. Esling, Cambridge 1999 (להלן: ספר יפ"א), עמ' 3-38.

3. מאתר זה אף ניתן להעתיק חינם את הגופנים (הפונטים) של לוח יפ"א למחשבי מקינטוש או למחשבי PC. באתר יש גם קישורים לרבים מתחומי ההתעניינות של העוסקים בפונטיקה בפרט ובבלשנות בכלל, ואפשר גם להאזין בו להקלטות של שפות שונות המתועתקות בתעתיק פונטי.

למעשה, אין בנמצא שיטת תעתיק שגלומות בה האפשרויות ששיטת יפ"א מציגה, ומכיוון שעדיין אפשר למצוא בישראל תעתיקים שלא בשיטת יפ"א, יש לשבח את החלטת מתכנני המאגר שבחרו להשתמש בשיטת התעתיק של יפ"א, המקובלת והנפוצה כיום בעולם המדעי. אלא שיש לנו כמה הערות מהותיות באשר לתעתיק המוצג במאמרו של יזרעאל (להלן: יזרעאל), הן באשר למה שיש בו והן באשר לחסר בו.

רצוי שהתעתיק והתמליל ילּוּ בהקלטה העובדה שלקורא אין גישה להקלטות שהתעתיק והתמליל התבססו עליהן היא מגרעת, כי אין בידי הקורא דרך לבדוק אם התעתיק והתמליל אכן מהימנים.⁴ מגרעת נוספת קשורה בזהות המתעתק והמתמלל. מהערה 1 במאמרו של יזרעאל משתמע כי מי שערך את התמלילים לא ראה בהכרח את התעתיק הפונטי, ולהפך – המתעתק לא הוא שתמלל.⁵ כל המנוסה בהכנת תמליל רק על סמך שמע אוזן מודע לעובדה שמתמללים שונים "שומעים" באופן שונה. דוגמה לכך, מהמאגר המודגם עצמו, מביא בן-טולילה בבקשו להדגים את הבעייתיות במלאכת התמלול: הוא מביא ארבע גרסות של תמליל להקלטה אחת.⁶ ידוע שתעתיק פונטי מסייע לתמלול ושתעתיק המגובה בניחות אקוסטי עדיף לעניין התמלול, כי הוא מסיר רבות מהעמימות. לפיכך מקובל כי מי שמתעתק הוא גם מי שמתמלל. אין זה מובן אפוא מדוע במאגר המודגם לא הוטל על מתעתק הדיבור גם לתמלל אותו (או לפחות להיות אחד מהמתמללים).

מכיוון שההקלטות לא צורפו למאמר, לא אתרכז בדבריי הבאים בכונותו או באי-נכונותו של התעתיק, אלא אעלה נקודות הקשורות לתעתיק עצמו.

תעתיק המאגר

במאמר הוסבר בסוגריים מרובעים נבחרו לסמן חפיפה כדיבור, וסוגריים צומדים לסמן תעתיק פונטי.⁷ כל העוסק בבלשנות יודע שתעתיק פונטי מסומן בסוגריים מרובעים דווקא, ולכן יש לתמוה מדוע בחר המחבר דרך שונה מהמקובל בעולם הבלשנות. בכל מקרה, שימוש בסוגריים מרובעים שלא לסימון של תעתיק פונטי מבלבל את הקורא.

4. בהערה 1 במאמרו של יזרעאל נאמר שהתעתיקים "הותקנו ועובדו [?] בידי יעל מימון" ושהגרסות הסופיות בתמליל ובהתעתיק "עובדו" על ידי יזרעאל.
5. שם, עמ' 290.
6. "בן-טולילה, "השיחה הטבעית כנושא למחקר", אקדם (ידיעון האקרמיה ללשון העברית) 22 (שבט תשס"ג), עמ' 6-7.
7. יזרעאל, עמ' 291.

אוריינטציה פונולוגית

התעתיק הפונטי במאמר הוא "צר למחצה"⁸ והדבר מקובל לצרכים כמו זה של מאגר לשון. ידוע שאפילו בתעתיק צר יש לפעמים כמה דרכים לסמן הגה אחד,⁹ ומכיוון שהתעתיק צריך לשרת מאגר שאפשר יהיה לשחזר בעזרתו אוצר מילים, למשל, רצוי שהתעתיק הפונטי ידמה ככל האפשר לתעתיק פונולוגי (כלומר התעתיק יהיה בעל אוריינטציה פונולוגית). לפי עיקרון זה יפה תועתקה המילה אז בסימן [z] ולא בסימן [s].¹⁰ אבל לפי אותה שיטה היה צריך לתעתק את מימוש הבי"ת במילה מבסטו¹¹ ב-[b] ולא בסימן [p], וכן את מימוש הבי"ת באוהב¹² ב-[v] ולא בסימן [f], את מימוש הדל"ת בדפוק¹³ ב-[d] ולא בסימן [t], את מימוש הדל"ת בחודשים¹⁴ ב-[dʒ] ולא בסימן [ð] ואת מימוש התי"ו במקלחת¹⁵ ב-[tʃ] (הסימן [t], ומתחתיו סימן לקוליות [ʃ], בצירוף הסימן לאופן חיתוך רווח יותר [l.]) ולא בסימן [ð].¹⁶

גם בתנועות יש לשאוף לאוריינטציה פונולוגית. בעברית המדוברת אמנם יש לאין סוף ביצועים פונטיים, המתפרסים על רוב שטחו של טרפו התנועות, אך אין-סוף התנועות הנה מתמיינן לחמש המשפחות של חמש התנועות הפונמיות בעברית.¹⁷ כמו בעיצורים, ראוי שסימן הבסיס בתעתיק הפונטי יהיה הסימן הפונולוגי והגיוון הפונטי האישי יסומן בסימנים דיאקריטיים. לפי זה יפה תעתקו את התנועה שאחרי מימוש הלמ"ד במילה לרמי בסימן [e].¹⁸ אם נניח ששמעו בהקלטה תנועה ממורכזת במילה לאסי,¹⁹ ראוי היה לסמנה [ë] ולא [ə], וכן במילה איזה²⁰ היה ראוי לסמן [ë] ולא [ə]. כן היה ראוי לסמן את זה שבאיזה²¹ בתנועה [ɛ] ולא ב-[e]. אף שמבחינה פונטית ייתכנו כמה אפשרויות סימון,²² ראוי שהאוריינטציה תהיה פונולוגית וכך תישמר גם העקיבות בשיטת התעתיק.

8. שם, עמ' 292 הערה 2.
9. ספר יפ"א, עמ' 30.
10. פעם בעמ' 313 ופעמיים בעמ' 314.
11. ירעאל, עמ' 313.
12. שם, עמ' 314.
13. שם, עמ' 313.
14. שם, שם.
15. שם, שם, בפעם הראשונה.
16. אגב, ללא עדות של הקלטה או של ספקטרוגרמה אני מתקשה להאמין שבהקשר זה הייתה גם הידמות בקוליות.
17. A. Laufer, "Hebrew", *Handbook of the International Phonetic Association (IPA)*, ed. J. Esling, Cambridge 1999.
18. ירעאל, עמ' 313.
19. שם, שם.
20. שם, עמ' 314.
21. שם, שם.
22. ספר יפ"א, עמ' 30.

עקיבות בתעתיק

בתעתיק שבעמ' 312 מופיע [ka'con] וזהו המקור לתמליל המילה רצון. גם בעמ' 306 מתועתקת הצד"י שבעצרי פעמיים בסימן [c]. השימוש בסימן זה מעורר תמיהה: בתעתיק יפ"א הסימן [c] מוגדר "סותם חיכי בלתי־קולי", הקיים, למשל, כפונמה בהונגרית והוא נשמע כסותם שבין [k] ל-[t]. אני מניח שנפלה כאן שגיאת תעתיק, אולי בהשפעתה של שיטת כתב שאינה של יפ"א, והכוונה המקורית הייתה לציין כאן "מחוכך מכתשי בלתי־קולי" שסימונו ביפ"א [ts] (אגב, בלוח יפ"א אין שורה מיוחדת למחוכך, כי אפשר להתייחס אליו כאל צירוף של סותם ואחריו חוכך שווה־בסיס־חיתוך. ואכן ביפ"א מסמנים מחוכך על ידי צירוף סימנים, וכשרוצים להדגיש שבשפה מסוימת מחוכך הוא פונמה אחת, יפ"א מאפשר להוסיף קשת מעל לסימנים או מתחתיהם: /tʃ/ או /tʃs/. מכל מקום, בתעתיק שלפנינו אין אחידות, כי שאר מימושי הצד"י בו אכן מסומנים [ts].²³

ל

דיוק בתעתיק

כאמור, ההקלטה לא צורפה לתעתיק, אך מניסיוני בתעתיק של עברית מדוברת ובניתוח אקוסטי שלה קשה לי להאמין בנכונות או בדיוק של כמה מסימני התעתיק שלפנינו. למשל, כבר שמעתי וראיתי אצל דוברים מסוימים ש-/n/ בסוף יחידת טעם (יוסבר להלן) מתקצרת וכי התנועה שלפניה מאונפפת, אך טרם מצאתי ש-/n/ בעברית נשמטת לגמרי. נכון ש"לא מצאתי" אינו ראיה, אך בכל זאת אני ממליץ לבדוק שוב את כל ההיקרויות של תופעה כזאת.²⁴

פעמים רבות מצאתי בניתוח אקוסטי של דיבור עברי טבעי, כי כשיש היקרות של שתי פונמות זהות בתוך קבוצת משמעות אחת, דוברי עברית מממשים רק פונמה אחת, ואפילו אותה אין הם מאריכים בדרך כלל (תופעה כזאת נצפתה גם בשפות אחרות). לכן הייתי ממליץ לבדוק שוב גם את התעתיק של רוח חבל ב-[xx].²⁵ לא מצאתי בעברית רהוטה שתי הברות מוטעמות סמוכות במילה אחת, ומשום כך סימון הטעמים על המילה מְבֹאֵסִית²⁶ נראה לי שגוי.

תעתיק בסיוע ניתוח אקוסטי

נעיר הערה כללית הקשורה לדיוק ולמהימנות של התעתיק: תעתיק פונטי על סמך שְמַע אוזן תלוי במיומנותו של המתעתק, ואפילו המתעתק מנוסה ומיומן, בסופו של דבר התעתיק הוא סובייקטיבי ולכן אין הוא נאמן. כיום, כשתכנות לניתוח דיבור

23. ראה אצל, אצלי וצאה שבעמ' 314.

24. למשל, פעמיים במילה נכון בעמ' 312; כאן ה-/n/ היא בתחילת הברה ובכל זאת נרשמה כאילו הושמטה!

25. יזרעאל, עמ' 313 והערה 8.

26. שם, עמ' 297.

זמינות, ראוי לכל מתעתק להסתייע בניתוח ספקטרוגרפי. שילוב של ניתוח על סמך שִׁמְע אוזן עם ניתוח אקוסטי מקובל כיום בעולם. הדבר רצוי מאוד, שכן הוא הופך את התעתיק הפונטי לאובייקטיבי יותר, ובסופו של דבר למדויק יותר.²⁷ לעניין זה נסתפק בדוגמה אחת: מתעתקים נוטים להשמיט הגאים הקצרים מסטנדרד מסוים, אך מתברר שכאשר הם מתכוננים בניתוח האקוסטי, הם רואים אפילו הגאים קצרים, ובהאזנה נוספת הם גם שומעים אותם.²⁸ בכל מקרה, ההתייחסות לתעתיק המגובה בניתוח אקוסטי זוכה לאמון רב של המשתמשים בתעתיק, ולכן רצוי שכל תעתיק פונטי יגובה בניתוח אקוסטי.

יחידות דיבור ברמה שמעל לרמה הפונולוגית

המחבר כותב שהלשון המדוברת מאורגנת ביחידות הנגנה, וטוען כי "קביעת גבולותיה של יחידת ההנגנה הן עניין מסובך הדורש מיומנות רבה". הוא מתלבט בשאלה "מה הם הסימנים הפורמליים ליחידה שאחריה אנו מצפים להמשך או אין אנו מצפים להמשך, וכיצד נסמן שאנו מפנים שאלה או בקשה לבן שיחנו". הוא ממשיך וקובע כי "שאלות אלה לא נתבררו עדיין בקשר לעברית [ההדגשה שלי]". המחבר מעלה "שאלה בסיסית נוספת" בעניין "מהות היחידות הקטנות מיחידת הנגנה, דהיינו היחידות המקבילות למילים".²⁹ על כל אלה יש להעיר כי שאלות כאלה נתבררו בשפות רבות וגם בעברית בת־זמננו. ראשית, במאמרי,³⁰ שזרעאל עצמו ציטט בעמ' 289, יש דיון על חלוקת הדיבור לקבוצות משמעות ולקבוצות נשימה, ומובא שם אף פירוט של הרמזים הפיזיים הגורמים לחלוקה, בצירוף דוגמות בעברית. שנית, במאמר מצוטט זה יש הפניות למקורות נוספים, ובהם ספרי בנושא הנגנה.³¹ אילו עיינו המחבר ושאר חברי הפרויקט

27. ראה למשל M. J. Ball and J. Rahilly, *Phonetics – The Science of Speech*, London 1999, pp. 139, 154; L. Lavoie, "Some Influences on the Realization of for and four in American English", *Journal of the International Phonetic Association* 32 (2002) 2, p. 189; J. E. Clark and C. Yallop, *An Introduction to Phonetics and Phonology*², Oxford 1995, פרק 7.

28. ראה למשל D. Binnepoorte and C. Cucchiarini, "Phonetic Transcription of Large Speech Corpora: How to Boost Efficiency Without Affecting Quality", *Proceedings of 15th ICPHS*, Barcelona 2003, p. 4; M. Wester, J. M. Kessens, C. Cucchiarini and H. Strik, "Obtaining Phonetic Transcriptions: A Comparison Between Expert Listeners and a Continuous Speech Recognizer", *Language and Speech* 44 (2001) 3, עמ' 390.

29. יזרעאל, עמ' 289.

30. א' לאופר, "הפסקה בדיבור רצוף ופיסוק", מחקרים בלשון העברית ובלשונות היהודים, מוגשים לשלמה מורג, בעריכת מ' בראשר, ירושלים תשנ"ו, עמ' 277–294.

31. א' לאופר, הנגנה, מפרסומי המכון למדעי היהדות, יב, ירושלים תשמ"ז (להלן: לאופר, הנגנה).

במקורות אלה, היו מוצאים שם שיטה לניתוח הנגנתי המתאימה לעברית המדוברת בת־זמננו.³² יתר על כן, אפילו לא היו מקבלים את השיטה המוצעת לעברית, אילו היו לומדים-מעיינים בספר הנגנה טוב כלשהו גם אם אינו מיוחד לשפה העברית, לא זו בלבד ששאלותיהם בעניין יחידות הדיבור היו נפתרות, אלא היו מגלים שבתעתיק המאגר שלהם הם השמיטו שני עניינים בסיסיים שבלעדיהם לא תיתכן הבנה של משמעות הדיבור. כוונתי לכך שגם בתמלילים וגם בתעתיקים שבמאמר לא סומנו לא ה"עיקריות" ולא ה"הנגנה".

כדי שנוכל לעמוד ביתר פירוט על חשיבותם ועל נחיצותם של סימני העיקריות וההנגנה, נקדים ונדגים כאן את מה שאנו מכנים "מודל התקשורת המילולית".

מודל התקשורת המילולית

הדיבור מורכב מרצף יחידות (אפשר לכנותן מקטעים, סגמנטים) המשתרשות זו אחר זו. מקובל שברמה הבסיסית המקטעים הם בגודל של יחידות הבניין הקטנות ביותר – פונמות. כלומר, ברמת הבסיס נוכל לראות בדיבור (הן זה המופק על ידי המדבר הן זה הנתפס על ידי המאזין) רצף לינארי של פונמות (עיצורים ותנועות). לפי מודל התקשורת שלנו גם הרמות הגבוהות יותר בנויות במתכונת דומה, מרצף לינארי של מקטעים. לדוגמה: רמת הדיבור שמעל הרמה הפונמית היא רמת ההברה, וגם אותה נוכל לראות כרצף לינארי של מקטעים, אך הפעם כל מקטע ארוך יותר וגודלו כגודל של הברה. ברמה שמעליה גודל המקטע הוא יחידת הטעם (לא נגדיר כאן את המונח, אך כדי לתת מושג־מה עליו נוכל לדמותו בפשטות ל'מילה' או ל'מבע', ומקובל גם המונח foot). ברמה הבאה גודל המקטע הוא קבוצת המשמעות (גם כאן לא נגדיר את המונח ולשם פשוט נוכל לדמותה ל'פסוקית'), וברמה שמעליה גודל המקטע הוא קבוצת הנשימה (בפשטות נוכל לדמותה ל'משפט').³³ ברמה שמעל קבוצת הנשימה גודל המקטע הוא קבוצת ההנגנה.

נוסף לתכונה המשותפת הזאת לכל הרמות, שכל אחת מהן מורכבת משרשור של מקטעים, נוכל לגלות עוד שתי תכונות משותפות לרמות השונות: (1) מקטע מורכב מאלמנטים שהם המקטעים שברמה שמתחתיו; (2) בכל מקטע יש אלמנט מרכזי

32. בחוברת התרגילים שסופחה לספרי הנגנה, תרגילים בהנגנה של עברית מדוברת,² ירושלים תשמ"ז, יש אלפי דוגמות לתרגול בעברית.

33. בפשטות נוכל לדמותה ל'פסקה'; להגדרה מלאה של המונח ראה לאופר, הנגנה, עמ' 64; עוד על "מודל התקשורת" ראה א' לאופר וא' טייטלבוים, "דיבור עברי מפי מחשב כחיקוי למודל תקשורת מילולית", דברי הקונגרס העולמי השנים עשר למדעי היהדות, חטיבה ד, ירושלים תשנ"ט; A. Laufer, Y. Sonnenblick and A. Teitelbaum, "Fundamental Frequency; Contour in Hebrew Synthesis by Rule – An Imitation of a Apeech Communication Model", *ICPhS-99, Proceedings of the 14th International Congress of Phonetic Sciences*, San Francisco 1999, עמ' 65–69.

אחד לפחות המלכד את שאר האלמנטים שבו. לדוגמה, ברמת ההברה נוכל להבחין (1) שמקטע ההברה מורכב מפונמות שהן האלמנטים שברמה שמתחתיו; (2) שבכל הברה יש גרעין המלכד את ההגאים באופן שהקשר ביניהם בתוך הקבוצה הדוק יותר מהקשר שבין הגאים השייכים להברות שונות: פונמה אחת (בדרך כלל תנועה) משמשת גרעין המלכד את שאר הפונמות שבהברה ליחידה אחת. באופן דומה, הרמה של יחידת הטעם מורכבת מהברות שהן האלמנטים של הרמה שמתחת ליחידת הטעם, ונוסף לכך יש ביחידת הטעם גרעין (ההברה המוטעמת) המלכד את הברות באופן שהקשר ביניהן בתוך הקבוצה הדוק יותר מהקשר בין הברות השייכות ליחידות טעם שונות.

להמחשת הדברים נביא להלן דוגמה ממשית של תקשורת מילולית בעברית. לפני כן נזכיר שבדיבור או בהאזנה איננו הוגים או שומעים פונמות שהן יחידות מופשטות. בדיבור אנו הוגים הגאים ממשיים, שהם נציגים של הפונמות אחרי שהושפעו מכל היחידות שברמות מעליהן, ואת הרמה הבסיסית הזאת נכנה "רמת האלופון". גם רמת האלופון היא שרשרת הגאים פיזיים שהדובר מפיך כדי להביע את רעיונו. רמה זו היא המגיעה באופן פיזי לאוזניו של המאזין, ומתוך ניתוח פיזי ולשוני של האלופונים מתפענחות רמות הדיבור השונות המאפשרות לו לתפוס את רעיון הדובר. במבצע כמו אני מבקש סליחה, אני לא מדבר עברית, החלוקה לרמות התקשורת בדיבור עשויה להיות כלהלן:³⁴

0. ברמת האלופון (בסימן ברמה זו אין כל משמעות לסימן הרווח):

[ʔani miʔakef səliΧa: ʔani loʊ meɔdabeɪ ʔiv.ʔitʰ]

1. ברמת הפונמה (בסימן ברמה זו אין כל משמעות לסימן הרווח):

/ʔa ni mevakef seliΧa ʔani lo medaber ʔivrit/

2. ברמת ההברה (הגבול בין המקטעים מסומן על ידי נקודות או על ידי רווחים):

/ʔa.ni me.va.kef se.li.Χa ʔa.ni lo me.da.ber ʔiv.rit/

3. ברמה של יחידת ההטעמה (המקטעים מסומנים על ידי קו תחתון וככל מקטע יש סימן טעם

אחד לפחות):

/ʔa'ni me.va'kef se.li'Χa ʔa'ni 'lo me.da'ber ʔiv'rit/

4. ברמה של יחידת קבוצת המשמעות (הגבול בין המקטעים מסומן על ידי קו מאונך |):

/ʔa'ni me.va'kef | se.li'Χa | ʔa'ni 'lo me.da'ber ʔiv'rit | /

5. ברמה של קבוצת הנשימה (הגבול בין המקטעים מסומן על ידי קו מאונך כפול – ||):

/ʔa'ni me.va'kef | se.li'Χa || ʔa'ni 'lo me.da'ber ʔiv'rit || /

34. רמת האלופון המוצגת היא תעתיק פונטי של דברי סגן נשיא ארצות הברית בשעתו, מר אל גור, בחגיגות יובל החמישים למדינת ישראל.

יש להדגיש שהמדבר משקע בדיבורו את רמות התקשורת השונות והן משפיעות זו על זו. מרמות הדיבור השונות נוצרת סינתזה שבסופה רק רמת האלופון המסונתזת היא היוצאת מפינו. אצל המאזין התהליך הפוך: לאוזניו מגיע גל קול שהוא תוצר של רמת האלופון הבסיסית של המדבר, ומהרמה הזאת הוא מסוגל לפענח את רמות התקשורת השונות.

לפי מודל התקשורת שלנו נוכל לתאר את רצף הדיבור כשרשרת של פונמות (הגאים), המתקבצות לשרשרת של הברות, המתקבצות לשרשרת של יחידות טעם (מילים), המתקבצות לשרשרת של קבוצות משמעות (או קבוצות מילים, פסוקיות), המתקבצות לשרשרת של קבוצות נשימה (או קבוצות מילים, משפטים), המתקבצות לשרשרת של קבוצות הנגנה (או שיח). בכל הרמות, למעט הרמה הבסיסית והרמה הגבוהה ביותר, נוכל לראות מבנה דומה: בכל מקטע יש מרכיב אחד לפחות המשמש גרעין היכול ללכד סביבו מרכיבים נוספים, וקבוצת כל המרכיבים היא המקטע של הרמה הנוכחית. כל אחת מרמות הדיבור תורמת לתקשורת ולפיכך יש לציינה בתעתיק הפונטי. בתמלילים ובתעתיקים שהוצגו במאמרו של יזרעאל לא סומנו שתי רמות דיבור חשובות הבאות לידי ביטוי בעיקריות ובהנגנה.

עיקריות וההבדל בינה ובין טעם

הרמה הגבוהה מרמת יחידת הטעם מכונה "קבוצת המשמעות" (sense group), והגרעין המלכד את הרמה הזאת מכונה "עיקריות" (prominence, ויש המכנים אותה accentuation). במבט הכתוב דן היה שוטר חרש כל אחת מהתיבות עשויה להיות עיקרית (לשמש גרעין של הקבוצה כולה), אך שינוי במיקום העיקריות גורר גם שינוי במשמעות של קבוצת המשמעות כולה. נסמן כאן את העיקריות בקו ניצב עילי כפול [1]. משמעו של המבט ['dan ha'ja [o'ter 'Xere]] עשוי להיות הנגדה של דן לעומת נפתלי; במבט ['dan ha'ja [o'ter 'Xere]] היה הוא גרעין הקבוצה, ומשמע המבט יכול להיות דן בעבר שוטר חרש, אבל עכשיו הוא איננו בתפקיד זה; במבט ['dan ha'ja [o'ter 'Xere]] חרש היא המילה העיקרית, ומודגש שדן היה שוטר חרש, כלומר שוטר סמוי, ולא שוטר שאינו שומע. המילה חרש הופיעה בשלוש דוגמות המבעים שלנו כשהיא מוטעמת מלעיל, אך רק בדוגמה האחרונה הייתה חרש גם המילה שיש בה טעם וגם המילה שהעיקריות חלה בה.

משתמע מכאן שגם העיקריות וגם הטעם הם גרעינים שתפקידם להבליט את להדגיש, אך אין הם שווים. ההבדל ביניהם הוא ברמה שהם מדגישים: טעם הוא הבלטה ברמה של יחידת טעם, ואילו עיקריות היא הבלטה ברמה של קבוצת המשמעות.³⁵

35. שאלה מעניינת היא כיצד אנו גורמים מבחינה פיזיולוגית או אקוסטית להבחנה בין טעם ובין עיקריות, אך אין זה המקום לדון בה.

סימון העיקריות הכרחי בתעתיק ובתמליל

עיקריות כשמה כן היא – ההבלטה של העניין העיקרי שהדובר רוצה להביע ביחידת המשמעות, ומיקומה הוא גורם התורם רבות למשמעות המבט כולו. שלוש הדוגמות שהבאנו לעיל המחישו ששינוי מקומה של העיקריות משנה את משמעות המבט כולו. כשם שברור לכול שאנו חייבים לסמן את מקום הטעם בעברית שכן בלעדיו לא נוכל להבחין למשל בין [Xereʃ] לבין [Xe'reʃ] או בין [bi'ra] לבין [bi'ra], כך גם לא יעלה על הדעת לתעתק מבע רצוף בלי לסמן את העיקריות. לפיכך ברור שאי אפשר לקבל תעתיק פונטי ואף לא תמליל, בלא שתסומן בהם העיקריות.

סימון העיקריות חשוב גם להבהרת שינויים פונטיים

בדיבור טבעי נמצא שיש נטייה להקפיד על חיתוך המילים העיקריות, וככל שההגאים רחוקים מהעיקריות יש נטייה "לחסוך" בחיתוכם. במילים שאינן "עיקריות", ובייחוד בהברות הלא-מוטעמות, רובנו נוטים לא להקפיד על חיתוך ההגאים, וכתוצאה מכך אנו מגלים בהן ביתר שאת את תופעות ההידמות וההשמטה.³⁶ בני אדם נבדלים במידת השימוש בהידמות ובהשמטות בדיבור, אך בכל מקרה העיקריות חייבת להיות מסומנת, כדי שיהיה אפשר לדעת אם ההשמטה היא תופעה טבעית הנגרמת מהריחוק מהעיקריות, או שמא היא תכונת דיבור אישית של הפרט הנבדק.

סימון הנגנה הוא דיווח על הכוונה הפנימית של המדבר

גם ההנגנה אינה מסומנת בתעתיק המובא במאמרו של יזרעאל, אף שיש לה משמעות. להמחשת חשיבות הסימון של ההנגנה נייער שוב במשפט שהבאנו לדוגמה, דן היה שוטר חרש. לצורך ההדגמה לא נשנה את מקומות הטעם ולא את מקום העיקריות והם יישארו קבועים, כמו בתעתיק זה: [dan ha'ja fo'ter 'Xereʃ]. כשנרצה להביע תמיהה נוכל להגות [dan ha'ja fo'ter 'Xereʃ], כשנרצה להביע שאנו מגלים סוד נגנן [dan ha'ja fo'ter 'Xereʃ], וכשנרצה לקבוע עובדה בלתי מעוררת נגנן [dan ha'ja fo'ter 'Xereʃ].³⁷

בעזרת ההנגנה הדובר מוסר לנו מידע חשוב על כוונתו הפנימית, על גישתו ועל הרגשתו. ההנגנה מסייעת בידנו לרדת לרוח הדברים שנאמרו ולקלוט מידע שאינו קיים בצירוף המילים כשלעצמן. לפיכך, אם מטרת מאגר של שפה מדוברת היא מחקר כללי של השפה, אי אפשר להסתפק בתעתיק ובתמליל ללא ציון ההנגנה, שכן

36. ראה פירוט בלאופר, הנגנה, עמ' 27 ואילך, וכן א' לאופר, פונטיקה ופונולוגיה, יחידות 4-5 בקורס מבוא לבלשנות, האוניברסיטה הפתוחה, תל-אביב תשנ"ב, עמ' 93.

37. הסתפקנו כאן בהדגמת שלוש הנגנות שונות שההבדל ביניהן הוא ב"נעימת הגרעין" בלבד. על הגדרת "נעימת גרעין" ועל הנגנות נוספות ועל דרכים לסימון בתמליל עברי ראה לאופר, הנגנה.

היא אחד הגורמים המרכזיים בהכנת הנשמע.³⁸ כיום אפשר להגיע ברשת האינטרנט לעשרות מאגרי דיבור בשפות שונות ולראות שיטות תעתיק של תופעות מקטעיות ועל-מקטעיות.³⁹

סיכום

כדי שיהיה אפשר להפיק תועלת ממאגר של עברית מדוברת, על ההקלטות להיות נגישות לקורא. האפשרויות הפשוטות כיום הן לקשר את ההקלטות לאתר ברשת האינטרנט או להעתיק אותן לתקליטור שיצורף לתעתיק ולתמליל. אין די בכך שהתעתיק יהיה בשיטת יפ"א. על המתעתק (שהוא גם המתמליל) להסתייע בניחוח אקוסטי כדי להגדיל את הדיוק ואת האובייקטיביות. כמו כן יש לכלול בתעתיק ובתמליל גם מרכיבים חיוניים בדיבור שהם מעל הדרגה של יחידת הטעם, כמו עיקריות והנגנה. אפשר אפוא לראות את התעתיק ואת התמליל המובאים במאמרו של יזרעאל רק כטיוטה מקדימה לתעתיק של עברית מדוברת.

38. ראה שם, ובייחוד בפרק "תפקידיה של ההנגנה".

39. למעוניינים נציין כי ברשת האינטרנט קיימים אתרים כלליים למאגרי לשון – <http://devoted.to/corpora>, <http://www.icp.inpg.fr/ELRA/>, <http://www ldc.upenn.edu> – lands.let.kun.nl/scratch/pubPERcat.php3?cat=3 – ומהם אפשר לפנות לעשרות מאגרי דיבור בשפות שונות ולקבל מידע על שיטות תעתיק של תופעות מקטעיות ותופעות על-מקטעיות במאגרי דיבור בשפות שונות.