

---

# סקירת ספרות בנושא מבחנים להערכת יכולת הבנת הנשמע בשפה שנייה

יניב גבאי-מילר

אוגוסט 2014



מרכז ארצי לבחינות ולהערכה (ע"ר)

NATIONAL INSTITUTE FOR TESTING & EVALUATION

المركز القطري للامتحانات والتقييم

מיסודן של האוניברסיטאות בישראל

**דוח מרכז 404**  
**ISBN:978-965-502-185-1**

**סקירת ספרות בנושא**  
**מבחנים להערכת יכולת הבנת הנשמע**  
**בשפה שנייה**

יניב גבאי-מילר

אוגוסט 2014



## תוכן העניינים

5	<b>הקדמה: מהי הבנת הנשמע</b>
6	<b>מבוא</b>
6	מדוע יש להעריך בנפרד את יכולת הבנת הנשמע
8	הגדרת המבנה
10	<b>סוגיות ודגשים בפיתוח מבחני הבנת הנשמע</b>
10	אופי הטקסטים ויצירתם
11	פורמט הקלט
13	כמות ההשמעות והזדמנות לקריאת השאלות
14	שפת הפריטים
15	סוגי מטלות
17	<b>סיכום</b>
19	<b>סקירת מבחנים בהבנת הנשמע מרחבי העולם</b>
19	Test of English as a Foreign Language (TOEFL)
21	International English Language Testing System (IELTS)
24	Pearson Test of English General (PTEG)
26	The European Language Certificates (TELC)
31	Test of English for International Communication (TOEIC)
33	Test Deutsch als Fremdsprache (TestDaF)
35	<b>מקורות לעיון נוסף</b>



## הקדמה: מהי הבנת הנשמע

הבנת הנשמע היא היכולת לעבד קלט אקוסטי – לעתים משולב בקלט חזותי – ולבנות ייצוג מנטלי נכון של תוכנו, ייצוג שיכול לשמש בסיס לתגובה (דבורה או כתובה). לא כאן המקום להרחיב על מודלים תאורטיים של תהליך הבנת הנשמע; נעמוד רק על כמה מאפיינים שלו החשובים לענייננו. הבנת הנשמע נעשית בכמה רמות לשוניות בעת ובעונה אחת, והמאזין משתמש בה בסוגים שונים של ידע לשוני, החל בידע פונולוגי וכלה בידע סוציולינגוויסטי. לדוגמה, לעתים אנו משתמשים בתחביר כדי לפענח מילה שהגייתה לא הייתה ברורה, או בסמנטיקה כדי להכריע בשאלות תחביריות. זאת ועוד, לא רק ידע לשוני משמש בפענוח הקלט: אנו משתמשים גם במידע ויזואלי, בידע כללי, בידע על נושא השיחה, בקונטקסט וכו'. הידע מאפשר לפסול אינטרפרטציות אפשריות של הקלט, וחשוב מזה, להשלים בדרך ההיסק פרטים שאינם נאמרים בטקסט במפורש – לא רק בעניין תוכן הטקסט עצמו, אלא גם בעניין מניעי הדובר, הפואנטה של הטקסט ועוד. בסך הכול יש בתהליך הבנת הנשמע אינטראקציה בין כמה מקורות מידע, והמאזין מסתמך עליהם במידה שונה לפי הצורך. אצל מאזינים שאינם דוברים ילידיים לעתים אין פרופורציה בין המשענות השונות, והם מנסים להבין את הנאמר בראש ובראשונה על סמך קשרים אסוציאטיביים בין מילות התוכן ועל סמך ידע העולם שלהם. הדבר עלול להיות בעוכריהם, בעיקר כאשר תוכן היחידה הרעיונית בלתי צפוי, או כאשר ההיסקים שהם עושים בהסתמך על ידע העולם שלהם שגויים.

בניית הייצוג המנטלי של הנאמר היא תהליך מתמשך, והייצוג המנטלי עצמו מצוי ביחסי גומלין עם הקלט האקוסטי: הוא משתנה עם קליטת מידע חדש, שנוסף על הקיים או מחייב שינויים בו, אך גם מספק קונטקסט להבנת הקלט. הקלט מעובד בכמה רמות טקסטואליות: על המאזין להבין את הלז של כל יחידה רעיונית כשלעצמה, אך גם את הדרכים שבהן היחידות הרעיוניות השונות מקושרות זו לזו. רוב הרעיונות החשובים בטקסט אינם כלואים ביחידה רעיונית אחת, אלא נבנים מכמה מבעים, מכמה יחידות רעיוניות שונות, ואף מכמה תורים בשיחה. לכן חשוב לא להסתפק בבחינת הבנה לוקלית בלבד.

### מדוע יש להעריך בנפרד את יכולת הבנת הנשמע

ארבע מיומנויות השפה הראשיות – הבנת הנקרא, הבנת הנשמע, הבעה בכתב והבעה בעל־פה – הן פנים שונים של אותו מבנה־על, השליטה בשפה, וכולן דורשות ידע פרוצדורלי בכל מרכיבי השפה. בין זוגות של מיומנויות מתקיימים קשרים הדוקים אף יותר, ובראש ובראשונה מתקיימים קשרים כאלה בין שתי מיומנויות־הנמען, הבנת הנשמע והבנת הנקרא, שמשותפת להן הדרישה הבסיסית לעיבוד של טקסט נתון ולהפקת ייצוג מנטלי של תוכנו. יש אפוא רגליים להנחה שבבקשו להעריך את מידת שליטתו של פלוני בשפה כלשהי, אנו רשאים לוותר על בחינה ישירה של יכולת הבנת הנשמע, ולהסתפק בבחינה של אחת או שתיים ממיומנויות השפה האחרות – ובפרט, של יכולת הבנת הנקרא – שייצגו את המבנה־על הכולל. ואולם, בין השפה הדבורה ובין השפה הכתובה יש גם הבדלים מהותיים, שבגינם ראוי להביא בחשבון בבחינה את יכולתו של הנבחן להתמודד עם שפה דבורה – ולמדוד את היכולת הזאת בנפרד מיכולתו להתמודד עם טקסט כתוב.

ההבדל הבולט ביותר בין שפה דבורה לשפה כתובה הוא באופן הקידוד והביצוע שלהן. השפה הדבורה מקודדת בצלילים, ולא באותיות, ולעתים קרובות הביצוע של הצלילים האלה חסר או לא־נורמטיבי. ביצוע חסר של צלילים אינו מפריע בדרך כלל להבנה אצל דוברים ילידיים: חלק ניכר מהמודיפיקציות הפונולוגיות הן סדירות, וממילא הנמען אינו זקוק לכל הקלט, אלא רק למספיק קלט כדי להפעיל את מקורות המידע האחרים שלו. אצל דוברים שאינם ילידיים, לעומת זאת, הגייה לא־נקייה היא גורם קושי של ממש. גם מבטא לא־מוכר עלול להקשות עליהם את ההבנה, ובמידה רבה יותר משיקשה אותה על דוברים ילידיים.

העובדה שהשפה הדבורה מקודדת בצלילים היא מקורו של הבדל קריטי אף יותר: הטקסט הדבור, שלא כטקסט הכתוב, הוא לינארי, והמאזין חייב לעקוב אחריו בזמן אמת. על פי רוב אין למאזין הזדמנות שנייה לשמוע את הטקסט הדבור, והוא אף אינו יכול להיוועץ במקרה הצורך בחלקים קודמים של הטקסט, שמהם נותר רק זיכרון – ולכן הוא מוכרח לעבד את הטקסט במהירות שבוחר הדובר, מהירות שהיא בדרך כלל גבוהה למדי. הבנת הנשמע חייבת אפוא להיות תהליך אוטומטי, ולא לערב תהליכים קוגניטיביים מודעים. אצל מאזינים שאינם דוברים ילידיים, לעתים התהליך עדיין איננו אוטומטי די הצורך – כך שכאשר קצב הדיבור גבוה, הם



אינם מספיקים להשלימו, ומתמקדים בהבנת הלקסיקון והדקדוק, תוך ויתור על העיבוד הרחב יותר של המשמעות; במהירות גבוהה אף יותר, הם לא יצליחו לקלוט גם את הלקסיקון והדקדוק, ויחמיצו חלקים גדולים מהטקסט. הרף הקריטי, שהחל בו הבנת הנשמע נפגעת, שונה אצל מאזינים שונים – וגורמי קושי נוספים על קצב הדיבור, למשל מבטא חזק, מנמיכים את הרף עוד יותר.

הטקסט הדבור שונה מהטקסט הכתוב לא רק בדרך ההעברה שלו, אלא גם במאפייניו הלשוניים. ראשית, יש בו משקל רב למאפיינים פרוזודיים, כמו אינטונציה, הממלאים שלל תפקידים, החל בהבעת היחס של הדובר לנאמר וכלה בבניית מבנה המסר של המשפט. הבנת הנשמע נשענת על המאפיינים הפרוזודיים לא פחות מאשר על התוכן המילולי; הפרוזודיה מכילה אלמנטים שמשלימים – ולעתים אף הופכים – את המשמעות המילולית של הנאמר. לעתים קרובות המאזין יכול לראות את הדובר, ואז הוא נעזר לא רק בקלט האקוסטי, אלא גם בסימנים לא-מילוליים, החל בתנועות קלות המסייעות לו בתהליכים כמו זיהוי פונמות וסגמנטציה, וכלה בשפת גוף ובהבעות פנים התורמים ישירות לבניית המסר של המוען. מאפיינים לשוניים אחרים המייחדים את השפה הדבורה נובעים מהלינאריות שלה ומהמהירות שהיא מופקת בה, המשפיעות לא רק על הנמען, אלא גם על הדובר: בדרך כלל אין בידו שהות לתכנן את הטקסט, והתוצאה לא-מאורגנת וחצי-אפויה. הטקסט הדבור אינו מורכב, כמו הטקסט הכתוב, ממשפטים בנויים לתלפיות המקושרים זה לזה בסמנים דקדוקיים פורמליים, אלא מיחידות רעיוניות קצרות יחסית, פשוטות מבחינה תחבירית ורזות מבחינת עומס המידע בהן, המשורשרות זו לזו בהתבסס בראש ובראשונה על קוהרנטיות רעיונית מיידית. בין היחידות – ולעתים קרובות גם בתוך – אפשר למצוא פאוזות, פילרים ("אה...") וחזרות שנועדו לקנות לדובר זמן, התחלות-שווא שהדובר מחליף מיד בטקסט אחר, תיקונים ותוספות-בדיעבד של מחשבה שנייה. לעתים קרובות הדובר משלב בדבריו בנדיבות יתירות וחזרתיות, המסייעות לנמען לעקוב אחר הנאמר. כמו-כן, השפה הדבורה שמרנית פחות מהשפה הכתובה, ונפוצים בה אלמנטים סובסטנדרטיים.

יש להדגיש שההבחנה בין שפה דבורה לשפה כתובה איננה דיכוטומית, אלא סקלרית: טקסט יכול להיות "יותר דיבורי" או "יותר כתוב" בכל מיני מובנים, החל באוצר המילים, עבור במידת הארגון שלו, וכלה בפונקציה שהוא משמש.

לסיכום, המיומנויות הנדרשות בהאזנה שונות מהמיומנויות הנדרשות בקריאה. אמנם הדרישות בתחומי אוצר המילים והניתוח התחבירי נמוכות יותר, אך כנגד זה המאזין נדרש לנקות ולארגן

בעצמו את הקלט: לזהות מילים אחרי שעברו מודיפיקציות פונולוגיות, להבין את האינטונציה של הדובר ולעתים גם את שפת הגוף שלו, לזהות אילו חלקים בקלט ראויים לעיבוד ומאילו יש להתעלם, לעמוד בעצמו על הקשרים בין היחידות הרעיוניות המרכיבות את הקלט – ולעשות כל זאת בזמן אמת, במהירות שקבע הדובר. אין תמה אפוא שהמתאם בין הישגי הנבחנים בפרקי הבנת הנשמע ובין הישגיהם בפרקי הבנת הנקרא באותן סוללות־בחינה – אף שהוא גבוה למדי, כמפורט להלן – אינו גבוה דיו כדי שחלק אחד יוכל למלא גם את מקום האחר (ליו וקוסטנזו 2013). המתאמים בין ההישגים בפרקי הבנת הנשמע ובין ההישגים בפרקים האחרים בסוללה נמוכים אף יותר. נראה אפוא כי המבנים המוערכים בפרקי הסוללה השונים קשורים זה לזה, אך רחוקים מלהיות זהים. מסקנה זו עולה בקנה אחד גם עם תוצאות ניתוח הגורמים של סווקי ושות' (2009) במבחני TOEFL, שמצא ארבעה גורמים מסדר ראשון (המתאמים לארבע מיומנויות השפה העיקריות) וגורם־על אחד.

מתאמים בין הישגי הנבחנים בהבנת הנשמע ובפרקים האחרים			
הבעה בעל־פה	הבעה בכתב	הבנת הנקרא	
0.739	0.612	0.789	TOEIC (ליו וקוסטנזו 2013)
0.654	0.643	0.735	IELTS (בוזורגיאן 2012)

שיקול נוסף שראוי להביא בחשבון הוא התוקף ההשלכתי של הבחינה: סביר להניח ששילוב פרק הבנת הנשמע בבחינה יעודד את המורים לעברית, ואת הנבחנים עצמם, להגדיל את ההשקעה בתרגול כישורים קומוניקטיביים.

#### הגדרת המבנה

יש להגדיר היטב ברמה התאורטית את המבנה שאנו מבקשים להעריך, ולדאוג לאופרציונליזציה הולמת שלו.

יכולת הבנת הנשמע איננה נקודה על ציר אחד, אלא נקודה במרחב רב־ממדי – ומספר הממדים בו

גבוה. בק (2001) מצטט כמה טקסונומיות שונות של תת-מיומנויות שנכללות כביכול במבנה התאורטי של הבנת הנשמע, ואף רשימות של תת-מיומנויות שמצא שהן מסבירות את מרבית השונות בין נבחנים בתחום זה: חלקן קשורות להצלחה בבחינות באופן כללי (היכולת להגדיר במדויק את המטלה ולשפוט אם מידע כלשהו רלוונטי לביצועה, היכולת לבנות תשובה במהירות וביעילות), חלקן קשורות להבנת שפה (למשל היכולת לזהות ולנצל יתירות, או היכולת לעבד מידע שמפוזר לאורך הטקסט), חלקן להבנה של שפות זרות (היכולת לעבד קונספטים של השפה הזרה שאין להם מקבילה מילולית בשפת האם), וחלקן ספציפיות להבנת הנשמע (בראש ובראשונה היכולת לסרוק ולעבד באופן אוטומטי טקסט מדובר במהירות). ואולם, גם הוא וגם לינץ' (2011) מזהירים מפני הסתמכות יתרה על טקסונומיות מסוג זה, ומדגישים שלרובן אין בסיס אמפירי, ואין לראות בהן אלא קווים מנחים שממחישים מה נחשב לחלק אינטגרלי מהבנת הנשמע, ויכולים לסייע בקביעת התת-מיומנויות שנרצה להדגיש בבחינה. טיילור וגרנפייה (2011) מוסיפים שקשה לעשות לרשימות כאלה אופרציונליזציה: קשה לדעת אם פריט כלשהו אכן נשען על תת-מיומנות מסוימת, ואם נשען, האם רק עליה – וייתכן גם שנבחנים שונים מפעילים תת-מיומנויות שונות לפתרון אותו פריט. בעיה זו עשויה להיפתר באמצעות נקיטת גישה אלטרנטיבית להגדרת המבנה, המבוססת לא על הגדרת התת-מיומנויות המרכיבות את יכולת הבנת הנשמע, אלא על משימות שאנו רוצים להעריך את יכולתו של הנבחן לבצע בהצלחה. ואולם, סימולציה של משימות אמתיות עלולה לגזול זמן רב, ולא זו בלבד, אלא שלא תמיד ברור מהן המשימות שיש לעשות להן סימולציה: אפילו במבחן כניסה ללימודים אקדמיים, המשימות שהנבחן יצטרך להתמודד אִתן בלימודיו משתנות לפי דיסציפלינה; ובמבחן כללי של שליטה בשפה עלינו להעריך את יכולת הבנת הנשמע של הנבחן בכל סיטואציית האזנה באשר היא, כך שאין מנוס מהגדרת המבנה בהתבסס על תת-מיומנויות.

ככל שמטרת המבחן ממוקדת יותר, כך קל יותר לבחור את התת-מיומנויות שצריכות לבוא בו לידי ביטוי. בכל מקרה רצוי לשים דגש על המאפיינים הייחודיים של הבנת הנשמע, שלא ייבחנו בשאר חלקי המבחן: היכולת לעבד ולהבין – באופן אוטומטי ובזמן אמת – טקסטים בעלי מאפיינים של שפה דבורה. על הטקסטים להיות מגוונים הן בתוכנם הן בסיטואציות ההאזנה שהם משקפים: חשוב לבחון גם את מיומנויות השיח של הנבחנים, ולכלול בבחינה גם מטלות הנשענות על ידע פרגמטי ועל הבנת השתמעויות שחורגות מהמשמעות המילולית של הנאמר.

## סוגיות ודגשים בפיתוח מבחני הבנת הנשמע

### אופי הטקסטים ויצירתם

יש מתח בין הרצון לשלוט בטקסט ולהתאימו לצורכי הבחינה ובין הצורך שהוא ישקף נאמנה מאפיינים של שפה דבורה, שלעתים קרובות איננה מאורגנת. קשה לשלב בטקסטים מתוסרטים היסוסים, פאוזות, טעויות ותיקונים באופן אמין, כך שהטקסטים אכן יישמעו כמו דיבור ספונטני. אמנם יש טקסטים שהם גם אותנטיים וגם מתוסרטים – לדוגמה, פרסומות, או מהדורות חדשות – אך קשה להצדיק ביסוס של הבחינה עליהם בלבד, שכן משתקף בהם רק חלק מהמבנה. עבור טקסטים אחרים יש לקבוע מה שיוויהמשקל הרצוי בין שני הצרכים, וכמה חופש פעולה יינתן לשחקנים: לעתים אין מנוס מכתובת הטקסט מראש וקריאתו בקול, אך בדרך כלל אפשר להסתפק בהנחיות (מפורטות יותר או פחות) לגבי התוכן, בלי להכתיב ניסוחים מדויקים. לעתים אפשר לוותר אפילו על הנחיות כאלה, ולהסתפק בקביעת נושא הטקסט: למשל, אפשר לבקש מסטודנטים להסביר רעיונות שונים מתחומי עיסוקם.

שימוש בשפה נעשה תמיד בתוך הקשר ולצורכי תקשורת, ומן הראוי שהדבר ישתקף בבחינה. יותר משחשוב לבחון את שליטתו של הנבחן בדקדוק הפורמלי של שפת היעד, חשוב להעריך כיצד הוא צפוי להשתמש בשפה בסיטואציות אמיתיות שבהן יזדקק לה. יש אפוא יתרון ביצירת טקסטים שידמו סיטואציות האזנה אמיתיות באופן ריאליסטי. עם זאת, יתרון זה עלול להיות חרב פיפיות: שלא ככללי השפה – שנשארים דומים בכל סיטואציה, ולפיכך בחינה שלהם צפויה לספק תמונה אמינה של השליטה הלשונית – סיטואציות האזנה שונות זו מזו, ואין ערובה שביצועי הנבחן בסיטואציה אחת ישקפו יכולת כללית שלו. נוסף על כך, הצורך בבניית קונטקסט שלם מאריך במידה ניכרת את הטקסטים, מכתובת הסתפקות במספר קטן יחסית של פריטי מבחן, ועלול למנוע גיוון מספיק שלהם, שיפצה על החיסרון הנ"ל. טקסטים ארוכים עלולים גם לעייף את הנבחנים (על אחת כמה וכמה אם הם לא מעניינים במיוחד), ויותר מדי קשיים בהבנה עלולים לגרום לנבחן לאבד את ידיו ורגליו ולוותר על המשך הטקסט. לכן רצוי שטקסטים ארוכים יהיו במבנה ברור יחסית, כך שמי שאיבד את הדובר יוכל לחזור לטקסט גם במהלכו. בכל מקרה, גם אם אי-אפשר לשלב בבחינה טקסטים באורך המשקף את סיטואציות ההאזנה האמיתיות שהנבחנים ייתקלו בהן, חשוב לבחון את יכולתם לבנות ייצוג מנטלי ברמת הטקסט השלם – גם אם הטקסט מורכב מכמה מקטעים קצרים.

חשוב לזכור שלטקסטים השפעה מכרעת על קושי הפריטים. יש אף רגליים לעמדה שעיקר הווריאציה בקושי צריכה להיות בטקסטים, ולא במטלות; לדוגמה, כדי לפתח פריטים קלים עדיף לצרף לטקסטים פשוטים, ואף ממושטים, מטלות שבוחנות באופן ראיסטי את עיבוד המידע אצל הנבחנים, ולא לצרף לטקסטים קשים, אפילו הם אותנטיים, מטלות שבוחנות רק הבנה שטחית שלהם. זאת מתוך התפישה שהאותנטיות של המשימה המוצבת בפני הנבחנים חשובה יותר מהאותנטיות של הטקסט: לב המבנה הוא היכולת להבין ולעבד באופן אוטומטי ובזמן אמת את הטקסט, ורצוי להיצמד לבחינת היכולת הזאת. על קושי הטקסט משפיעים שלל גורמים: כמובן גורמים הקשורים לתוכן הטקסט (מידת ההיכרות של הנבחנים עם הנושא, מורכבות הרעיונות המובעים בטקסט) ולשונו (מורכבות התחביר, אוצר המילים, מידת היתירות וכו'), וכן גורמים הקשורים לדרך ההעברה שלו, כמו המבטא וקצב הדיבור. שוהמי וענבר (1991) מצאו קשר הפוך בין מידת הדיבוריות של טקסט ובין רמת הקושי שלו: ידיעות חדשותיות היו קשות להבנה יותר מהרצאות קצרות לפני קהל, וההרצאות היו קשות יותר מדיאלוגים בין מומחה לאדם הנועץ בו. הטקסטים שהשתמשו בהם לא נבדלו זה מזה בתוכנם, אלא במידת הדיבוריות של לשונם: המידה שבה הם היו מתוסרטים, כמות החזרות, היתירות, הפרעות והפאוזות, מידת האינטראקציה בין הדובר לנמען, מורכבות המשפטים והתקינות הלשונית. תוצאה זו מחזקת את ההשערה שתוקף הבחינה יצא נשכר משילוב טקסטים במגוון רמות של דיבוריות. אשר לקצב הדיבור, בלאו (1990) מצאה שהאטתו מסייעת בעיקר לנבחנים ברמות היכולת הנמוכות – ושהכנסה של פאוזות בגבולות תחביריים ורעיוניים מסייעת לנבחנים אף יותר ממנה. הכנסת פאוזות במקומות אקראיים בטקסט דווקא מקשה על הבנתו.

### **פורמט הקלט**

כאמור לעיל, בסיטואציות האזנה אמתיות, לעתים קרובות המאזין נעזר לא רק בקלט האקוסטי, אלא גם בסימנים ויזואליים. שימוש בווידאו יכול אפוא לשפר את האותנטיות של חלק מהפריטים. וידאו עשוי גם לסייע לנבחנים להבחין בין הדוברים השונים ולזהות את תפקידיהם ואת היחסים ביניהם, וכן לספק להם מידע על הסיטואציה וההקשר שהדברים נאמרים בהם. המידע מסייע לנבחנים למקם את עצמם תודעתית בסיטואציית ההאזנה המתאימה: אם יראו שהם עומדים לשמוע שיחה בין רופא לחולה, לדוגמה, יצפו לביטויים ולנוסחאות שיחה אופייניים לשיחות כאלה – ממש כשם שהיו מצפים להם אילו עמדו להשתתף בשיחה כזו בעצמם. ואכן,

וגנר (2010) מצא כי הישגיהם של נבחנים בבחינות הבנת הנשמע שצפו בווידאו היו גבוהים ב-6.5% בממוצע מהישגי הנבחנים שהאזינו לאותם הטקסטים באודיו בלבד, ולינץ' מזכיר כמה מחקרים שעולה מהם ששימוש בכמה ערוציתקשורת במקביל – דיבור, כתב, תמונה ושפתגוף – משפר את הבנת הנשמע אצל מאזינים מתחילים ומתקדמים כאחד. עם זאת, הווידאו אינו מסייע לכל הנבחנים באותה מידה: יש שונות גדולה בין נבחנים בכמות הקשב שהם מפנים לווידאו (וגנר 2007), נבחנים שונים נבדלים ביכולת שלהם לפרש סימנים ויזואליים עדינים, ונבחנים מרקע תרבותי שונה הווידאו עלול אפילו להטעות. וגנר סבור שמשתנים אלו הם חלק ממבנה הבנת הנשמע – שהרי במרבית סיטואציות ההאזנה המאזין יכול לראות את הדובר – ומזהיר כי ויתור על בחינת יכולתם של הנבחנים להיעזר בסימנים ויזואליים עלול לגרום לפגיעה בתוקף הבחינה בשל ייצוגיחסר של המבנה; לינץ' טוען שלנוכח השינויים בסביבת הלמידה האקדמית "נעשה קשה יותר ויותר להצדיק הערכה של הבנתנשמע אקדמית בהתבסס על קלט אקוסטי בלבד [...]. (מחקרים) מדגישים את הצורך לנצל את סוג הקלט המשלב בין אודיו לווידאו שהסטודנטים ייתקלו בו בחייהם האקדמיים, בין אם בסביבת למידה מסורתית ובין אם באינטרנט או בלמידה מרחוק". אחרים, ובהם בק, אינם רואים ביכולת לפרש סימנים ויזואליים חלק מן היכולת הלשונית, ומזהירים בהתאם לכך כי דווקא שילוב וידאו בבחינה עלול לפגוע בתוקף שלה, בשל משתנים מתערבים שאינם רלוונטיים למבנה. גם לפי וגנר (2010), עדיין אין די ראיות אמפיריות התומכות בהשערה ששימוש בווידאו משפר את התוקף של מבחני הבנת הנשמע.

חיסרון ודאי של השימוש בווידאו הוא העלות: הפקת פריטים המשתמשים בווידאו יקרה ומסובכת יותר מהפקת פריטים המשתמשים באודיו בלבד, והאמור נכון גם בנוגע להעברתם. לפחות חלק מיתרונותיו של השימוש בווידאו אפשר לנסות להשיג גם בדרכים אחרות: בין כך ובין כך רצוי שהקולות המוקלטים יהיו שונים מאוד זה מזה, ואפשר גם להקדים לטקסט המושמע כותרת ("אצל הרופא") או הסבר קצר. עוד פתרוןביניים אפשרי הוא לצרף לפריטים תמונות סטילס, שיכולות לסייע בזיהוי ההקשר והדוברים. יש להדגיש שלא כל תמונת אילוסטרציה תמלא את התפקידים האלה: כפי שמעירה גינטר (2001), מקום התרחשותן של שיחות, הנראה בתמונה, קשור רק לעתים רחוקות לתוכן, ותצלום גנרי של מרצה ישוב ליד שולחן אינו מועיל לנבחנים לעקוב אחר ההרצאה. על עזרים ויזואליים, בין אם תמונות סטילס ובין אם וידאו, להשלים את הטקסט הדבור ולתמוך בו – אם כי בכל מקרה יש לוודא, כמובן, שהשלמת המטלה תלויה בהבנת הטקסט.

למותר לציין שאיכות הקלט צריכה להיות מעולה : בראש ובראשונה יש להקפיד על אקוסטיקה טובה באתר הבחינה, על מזעור רעש הרקע ועל איכות הציוד וההקלטה. רצוי אף להשתמש באזניות באיכות טובה.

### **כמות ההשמעות והזדמנות לקריאת השאלות**

החד-פעמיות היא אחת התכונות המרכזיות של שפה דבורה : בדרך כלל יש למאזין רק הזדמנות אחת לשמוע את הטקסט, וגם אם מתאפשר לו לבקש מהדובר לחזור על דבריו, כמעט תמיד הדובר ישנה את הניסוח, הדגשים, או לכל הפחות האינטונציה. סביר להניח גם שעיבוד של טקסט בשמיעה חוזרת מפעיל מיומנויות אחרות מעיבוד שלו בשמיעה יחידה. מאידך גיסא, יש לזכור שפריטי מבחן דורשים לעתים קרובות הבנה מדויקת יותר משנדרשת בסיטואציות האזנה אמתיות – וכשנדרשת בסיטואציות כאלה הבנה מדויקת, לא פעם יש לנו הזדמנות לשאול ולבקש הבהרות. כמרכן, השמעה כפולה של הטקסט מקטינה את הפגיעה בנבחנים שיחמיצו חלקים ממנו בשל רעש רקע. יש אפוא טיעונים טובים לכאן ולכאן, וההכרעה ביניהם תלויה במטרות הבחינה. עקרונית, נראה שהן האוטנטיות הן תוקף המבנה מכתיבים השמעה אחת בלבד של הטקסטים בבחינה – אלא במקרים יוצאי-דופן, שבהם גם בסיטואציית ההאזנה האמתית יש חזרה, כמו כריזה בתחנת רכבת או הודעות במשיבון. כדי להתגבר חלקית על בעיית הדיוק אפשר לשלב בטקסט חזרות ופרפרזות, שכמותן יש רבות גם בטקסטים אמתיים. בדיאלוגים אפשר לשלב גם בקשות מפורשות לחזרה או להבהרה.

בסיטואציות שיחה אמתיות, הצדדים משתמשים בשפה כדי להשיג מטרה תקשורתית כלשהי, והמאזין בדרך כלל יודע איזה מידע הוא מחפש. מומלץ לחקות מצב זה באמצעות מתן הזדמנות לנבחנים לקרוא את השאלות לפני ההאזנה לטקסט. הידיעה מה עליהם למצוא מאפשרת לנבחנים לבחור באסטרטגיית האזנה מתאימה ולהתמקד במידע החשוב להם. יוצאים מכלל זה פריטים הנוגעים לעניינים מהותיים או גלובליים, כמו המסר המרכזי של דובר כלשהו, או הרעיון המרכזי בהרצאה.

אפשר להגביר את העומס הקוגניטיבי על הנבחנים אם מבקשים לפתור את השאלות תוך כדי ההאזנה לטקסט, ולא מקצים זמן מיוחד לפתרון אחריה. גם כאן רצוי לשמור על אותנטיות: פתרון השאלות תוך כדי האזנה עשוי לדמות כתיבת רשמים בהרצאות, ואפילו סיטואציות

חברתיות, אך פחות מתאים להאזנה לריאיון רדיופוני, למשל. בכל מקרה חשוב לשמור על ריווח בין השאלות, ולא להשמיע בעת פתרון שאלה מידע חיוני לפתרון השאלות הבאות אחריה.

## שפת הפריטים

בבחינת הבנת הנשמע, שלא כבבחינות כתיבה או דיבור, הנבחן אינו מפיק תוצר ישיר שאפשר לבחון – התוצר הוא ייצוג מנטלי בדעתו, והוא מתוּוֹךְ לנו באמצעות פריטי המבחן. זוהי פרצה הקוראת לגנב, כלומר – למשתנים מתערבים. לדוגמה, בשאלות בררה על הנבחן לקרוא את כל אפשרויות התשובה, ויש התערבות של יכולת הקריאה שלו – ובמטלות שהנבחן צריך לנסח בהן תשובה בעצמו, מתערבת יכולת ההתנסחות בכתב או בעל־פה.

דרך אחת להפחית את השונות הלא־רלוונטית היא לצמצם את השונות בין הנבחנים במשתנים המתערבים. לדוגמה, סביר להניח שהשונות בין הנבחנים ביכולת הבנת הנקרא שלהם תהיה נמוכה יותר בשפת האם מאשר בשפת היעד. יש טעם אפוא לכתוב את פריטי המבחן בשפת האם של הנבחנים, אם הדבר אפשרי (אם כי יש להיזהר שהפריטים עצמם לא יספקו מידע שיסייע להבנת הנשמע). עם זאת, יש לציין שפיליפי (2012) מצאה מתאם גבוה בין הישגי נבחנים בפריטים שנכתבו בשפת האם שלהם ובין הישגיהם בפריטים שנכתבו בשפת היעד, כך שהתלות בהבנת הנקרא כנראה נסבלת. על פי רוב, השאלות בשפת היעד היו מעט קשות יותר, במיוחד אם הן כללו אוצר מילים לא־טרוויאלי – אך בחלק מהמבחנים שנבדקו הפער היה זעום, ופריטים שעסקו בנושא שחזר כמה פעמים בטקסט, או שהייתה חזרה לקסיקלית בין הטקסט ובין התשובה הנכונה בהם, נטו אף להיות קלים יותר כשנכתבו בשפת היעד. בכל מקרה כדאי להקצות בנדיבות יחסית את הזמן לקריאת השאלות, ולנסח את המטלה בפשטות המרבית האפשרית – ככל שהזמן המוקצב לקריאה קצר יותר, וככל שלשונו של הפריט מורכבת יותר, כן גדלה התלות בהבנת הנקרא. הדבר נכון שבעתיים עבור מבחנים בשפה העברית, שמערכת הכתב הזרה בהם כופה על הנבחנים שלב נוסף של עיבוד. באופן דומה, בפריטים הדורשים מהנבחנים להפיק תשובה בעצמם, רצוי להתעלם משגיאות כתיב, דקדוק וכדומה, כל עוד הן אינן מפריעות לוודא שהנבחנים הבינו את ששמעו.



## סוגי מטלות

על הפריטים להעריך את הידע והמיומנויות הנכללים במבנה, ולהיות מאוזנים במיומנויות שהם בוחנים וברמת הקושי שלהם.

המטלות הפשוטות ביותר הן מטלות דיסקרטיות: מטלות המבודדות אלמנט לשוני אחד ובחנות אותו בנפרד. דוגמה למטלה כזו היא הבחנה בין פונמות: לנבחן מושמע מבע קצר, ללא קונטקסט ("בפינה הוצב שלטי"), ועליו לבחור לגבי אחת המילים במבע איזו גרסה מהמוצעות לו ("שלט / שלד") היא זו ששמע. מטלות כאלה בדרך כלל אינן אותנטיות: כאמור בהקדמה, הבנת הנשמע נבנית תוך אינטראקציה מתמדת בין מקורות שונים, ואינה תלויה דווקא בהתמודדות עם אלמנטים ספציפיים בקלט. רצוי אפוא להשתמש במטלות דיסקרטיות רק אם יש הצדקה יוצאת-דופן לבידוד אלמנט אחד (למשל אוכלוסיית נבחנים ברמה נמוכה).

על פי רוב, נרצה לבחון הבנה ברמת היחידה הרעיונית לכל הפחות. דרכים טובות לעשות זאת הן מטלות שיפוט, התאמה, העברה וחזרה. במטלות שיפוט על הנבחן להעריך תוכן של מבעים, לדוגמה אם הם אמתיים, או שקריים. כדי למזער את התלות בידע עולם, אפשר לקשר את המבעים לתמונות; לחלופין, אפשר להשמיע לנבחן שני משפטים, ולשאול אם הם שקולים. מטלות כאלה מחייבות להבין את משמעות הנאמר בשלמותו. מטלות התאמה דומות להן במידת-מה: על הנבחן למצוא מבין כמה אפשרויות – למשל, כמה איורים – את האפשרות המתאימה לטקסט ששמע (או להפך). במטלות העברה על הנבחן לתרגם את הנאמר לפורמט אחר: לצייר בית ששמע תיאור שלו, לשרטט מסלול על-פי הכוונה וכו'. במטלות חזרה, הנבחן מתבקש לחזור על הנאמר, בכתב או בעל-פה. המיומנויות הנדרשות לכך תלויות באורך הטקסט: אם הוא קצרצר ופשוט, החזרה אינה אלא תרגיל בתעתיק, ודורשת רק זיהוי של המילים; מבעים ארוכים יותר עשויים לדרוש גם שימוש בזיכרון עבודה, חלוקה תחבירית ושחזור של מילים לפי הקשר. לפיכך, מקובל לפתוח מבחני חזרה במקטעים קצרים יחסית של טקסט, ולעבור בהדרגה למקטעים ארוכים יותר – מה שמאפשר להבחין בין נבחנים מרמות יכולת שונות. כיוון שייצוגים מנטליים מבוססים על הלזז של הקלט, ולא על הניסוח המדויק שלו, רצוי להימנע מדרישה לחזרה מדויקת על הנאמר, מילה במילה.

מטלות מהסוגים שתוארו עד כה הן קצרות ומהירות יחסית, והדבר מאפשר לשלב במבחן פריטים רבים מסוגים אלו. מאידך גיסא, הן מתמקדות בהבנה מילולית של הנאמר, ללא היסקים והשתמעויות. זאת ועוד, בדרך כלל הן בוחנות הבנה מקומית בלבד, ברמת היחידה הרעיונית,

וחלק חשוב מהמבנה הוא יכולת הבנה ועיבוד ברמת הטקסט השלם. דרך אחת לבחון יכולת זו היא לבקש מהנבחנים לכתוב סיכום של הטקסט. למטלות כאלה תוקף נראה גבוה (פאוארס 1986), אך הן דורשות חיבור מחוון ונקידה ארוכה, וחשופות לשונות לא־רלוונטיות הקשורה ביכולת ההבעה בכתב של הנבחנים. מטלות אחרות הדורשות הבנה ברמת הטקסט הן מטלות של השלמת פערים. דוגמה טובה למטלה כזו היא קלוז: הנבחנים שומעים פעמיים טקסט שכמה מהמילים בו הוחלפו בצפצוף – פעם אחת את כל הטקסט ברציפות, ואז שוב, עם הפסקה קצרה בגבול התחבירי הראשון אחרי כל צפצוף, שבמהלכה הם כותבים את המילה החסרה. הקלט החסר מושלם לפי ההקשר, הן הצר הן הרחב, כך שהנבחנים נדרשים להבין את כל הנאמר, גם החלקים שהם אינם משתמשים בהם ישירות. דוגמה אחרת היא השלמה של פערים בטקסט כתוב, המבוססת על המידע בטקסט הנשמע: למשל, הנבחנים מקבלים מראש סיכום כתוב של הטקסט הנשמע שמילות התוכן החשובות בו חסרות, ומשלימים את החסר בהסתמך על הטקסט הנשמע. (כמובן, יש לוודא שאי־אפשר להשלים את החסר בלי הטקסט הנשמע).

הדרך הישירה ביותר לבחון את הבנת הנשמע היא באמצעות שאלות ישירות על הנאמר. חשוב לשלב בין שאלות גלובליות, הדורשות סינתזה של מידע או הסקת מסקנות, לשאלות מקומיות, הדורשות איתור פרטים רלוונטיים או הבנת מבעים ספציפיים. רצוי להתמקד במה שיש בטקסט – במפורש או במשתמע – ולא במה שאין בו; בשאלות נכון / לא־נכון, לדוגמה, הנבחנים לא יכולים לחזור לטקסט ולבדוק את המשפטים, וגם לא יכולים לוודא אם מידע כלשהו נתון בטקסט או לא.

רצוי, כאמור, לצמצם את התלות בכישורי הקריאה והכתיבה של הנבחנים. בשאלות סגורות, רצוי לשמור על אפשרויות תשובה קצרות ותמציתיות, ואולי לתת אשכולות של כמה שאלות עם אותן אפשרויות תשובה (למשל, לתת רשימה של טענות, ולבקש מהנבחנים לסמן עבור כל טענה את הדובר או הדוברים שהיא עולה בקנה אחד עם דבריהם). בשאלות פתוחות, אפשר להגביל מאוד את אורך התשובה המבוקשת. הדבר מסייע הן בקביעת רשימת התשובות הקבילות, הן בשליטה בזמנים שהנבחנים מקדישים לשאלה (שכן הדרישה לענות בקיצור נמרץ פוטרת את הנבחנים מהצורך להקדיש זמן רב לניסוח תשובותיהם). עם זאת, שאלות כאלה מתאימות בעיקר לבחינת הבנת מידע שנמסר בטקסט במפורש, ונוטות אפוא להיות נקודתיות ושטחיות.

בכל מקרה על השאלות לוודא שהנבחן עיבד את הנאמר והבין אותו. רצוי שהתשובה לא תהיה ציטוט ישיר מהטקסט הנשמע – ומובן שעל הפריטים להיות תלויים בטקסט, כך שאי־אפשר לפתור אותם בלעדיו.

הבנת הנשמע היא היכולת לייצר ייצוג מנטלי נכון של תוכנו של קלט אקוסטי כלשהו. בתהליך זה המאזין נעזר בכמה מקורות מידע, לשוניים ולא-לשוניים, המצויים באינטראקציה מתמדת זה עם זה, ובמגוון רחב של תת-מיומנויות. חלק מהתת-מיומנויות משותפות גם לשאר מיומנויות השפה, ובראש ובראשונה למיומנות-הנמען האחרת, הבנת הנקרא; ברם רבות מהן ייחודיות להבנת הנשמע, בשל ההבדלים הניכרים בין שפה כתובה ובין שפה דבורה. ראוי שבחינה בהבנת הנשמע תתמקד בפנים הייחודיים האלה, כלומר ביכולת לעבד ולהבין בזמן אמת טקסטים בעלי מאפיינים של שפה דבורה.

במהלך פיתוח הבחינה יש להביא בחשבון שלל סוגיות, ובהן:

- **אופי הטקסטים:** רצוי לשלב בבחינה טקסטים מגוונים בנושאים, ברמת הקושי שלהם ובמידת הדיבוריות של לשונם. חשוב לבחון את יכולתם של הנבחנים להתמודד עם טקסטים שלמים, אך רצוי לא להעמיס עליהם מידע רב מדי ללא עזרי התמצאות.
- **פורמט הקלט:** שימוש בוידאו עשוי לשפר את האותנטיות של מרבית הפריטים. עם זאת, הוא יקר, ואין קונצנזוס באשר לשאלה האם תרומתו לשוונות בין הנבחנים רלוונטית או לא. חלק מהיתרונות שבשימוש בוידאו אפשר להשיג גם באמצעות שימוש בתמונות סטילס.
- **כמות ההשמעות:** השמעה חוזרת של טקסט עשויה לפצות על הפגיעה המסוימת באותנטיות שנגרמת לעתים מהיעדר היכולת של המאזין לתת פידבק לדובר, ומקטינה את הסיכוי שנבחנים יחמיצו את הטקסט בשל רעש רקע. עם זאת, אפשר להשיג את היתרונות האלה גם באמצעים טקסטואליים – וההשמעה החוזרת עלולה לפגוע הן באותנטיות של הפריט, הן בתוקף המבנה שלו.
- **זמן קריאת השאלות ופתרון:** מתן הזדמנות לקריאת הפריטים לפני השמעת הטקסט הופך את ההאזנה למוכוונת-מטרה, ומתאים אפוא למטלות המדמות סיטואציות שזהו סוג ההאזנה הנדרש בהן. גם ההחלטה אם להקצות זמן מיוחד לפתרון השאלות ראוי שתעלה בקנה אחד עם הסיטואציה שהמטלה מדמה.

- **שפת הפריטים** : פריטים הכתובים בשפת הבחינה עלולים לערב משתנים זרים למבנה, כמו יכולת הבנת הנקרא או הביטוי בכתב. אם אין אפשרות לכתוב את הפריטים בשפת האם של הנבחנים, יש לנסח את הפריטים בפשטות, להקצות לקריאתם זמן רב, ובפריטים פתוחים – להתעלם מתכונות לא רלוונטיות של התשובה.
- **סוגי המטלות** : יש לשלב בבחינה מטלות הדורשות מגוון מיומנויות, שיהיו מדגם מייצג של התת-מיומנויות הנכללות במבנה. מטלות המדמות סיטואציות האזנה אמתיות, גם אם בזעיר אנפין, הן אותנטיות במיוחד, אך עלולות להגביל את מספר הפריטים בבחינה.

## סקירת מבחנים בהבנת הנשמע מרחבי העולם

### Test of English as a Foreign Language (TOEFL)

מבחן TOEFL מיועד בראש ובראשונה למעוניינים ללמוד באוניברסיטאות בשפה האנגלית. מרבית הנבחנים בו נבחנים בגרסה האינטרנטית האדפטיבית שלו, המורכבת מארבעה פרקים: הבנת הנקרא, הבנת הנשמע, הבעה בעלפה והבעה בכתב. בפרק הבנת הנשמע 6–9 טקסטים באורך 3–5 דקות כ"א, המדמים סיטואציות האזנה אופייניות לעולם האקדמי: 2–3 שיחות בין סטודנט לאיש סגל כלשהו, 4–6 הרצאות קצרות, לעתים עם הערות מהקהל. בסך הכול מוקצות לפרק 60–90 דקות. לכל שיחה מצורפות חמש שאלות, ולכל הרצאה – שש. השאלות הן שאלות בררה וכתובות באנגלית.

כל טקסט מתחיל בהצגה כללית (למשל בנקיבה בתפקידי הדוברים, או בתחום הדעת של הרצאה) ומושמע פעם אחת בלבד, והנבחנים אינם רואים את השאלות לפני ההשמעה. עם זאת, הם רשאים לרשום לעצמם הערות בזמן ההאזנה. מרבית השאלות נוגעות לטקסט בכללותו (למשל, מהו הנושא העיקרי שנידון בו) או לנקודות מפתח בו, שיש יסוד סביר להניח שהנבחנים רשמו בהערותיהם; הפורמט משקף היטב את הבנת הנשמע הנדרשת בהרצאות אקדמיות, אם כי עלול להחמיץ את הנדרש בסיטואציות יום-יומיות מוכוונות-מטרה. אחדות מהשאלות נוגעות להשתמעויות מניסוחים ספציפיים בטקסט, ולפניהן הניסוחים האלה מושמעים שוב. לדוגמה:

(Female student) I'm sorry I had to miss practice, though. I feel bad about that.

(Male coach) Family's very important.

What does the man mean when he says: "Family's very important"?

- a. He hopes the woman's family is doing well.
- b. He would like to meet the woman's family.
- c. The woman should spend more time with her family.
- d. The woman had a good reason for missing practice.

בהתאם ליעדי המבחן, השאלות מתמקדות בהבנות גבוהות יחסית: הבנה של רעיונות מרכזיים, של יחסים בין רעיונות בטקסט, של השתמעויות ושל הלך הרוח של הדובר.

חלק מהשאלות דורשות קריאה רבה יחסית (אם כי לא יותר משורה לכל אפשרות תשובה). הדבר

עולה בקנה אחד עם המדיניות הכללית במבחן, לשלב בפריטים כמה מיומנויות: לדוגמה, בחלק מהפריטים בפרק הדיבור הנבחן נדרש להגיב בעלפה על דברים שקרא או ששמע. השילוב בין המיומנויות השונות בוודאי אותנטי ומשקף את שיידרש מהנבחנים באקדמיה, אך עלול לפגוע בתוקף של כל פרק בנפרד.

הטקסטים עשירים יחסית במאפיינים אורליים, הן ברמה המקומית הן ברמה הגלובלית. לדוגמה, קטע מתוך הרצאה על הספר "אבן הירח":

Now, Cuff, this extraordinarily important character... well, let me try to give you a sense of who Sergeant Cuff is, by first describing the regular police. And this is the dynamic that you're going to see throughout the history of the detective novel, where you have the regular cops - who are well-meaning, but officious and bumblingly inept - and they are countered by a figure who's eccentric, analytical, brilliant, and... and able to solve the crime. So, first the regular police get called in to solve the mystery - Um, in this case, detective, uh, Superintendent Seegrave. When Superintendent Seegrave comes in, he orders his minions around, they bumble, and they actually make a mess of the investigation, which you'll see repeated - um, you'll see this pattern repeated, particularly in the Sherlock Holmes stories of a few years later where, uh, Inspector Lestrade, this well-meaning idiot, is always countered, uh, by Sherlock Holmes, who's a genius. So, now Cuff arrives. Cuff is the man who's coming to solve the mystery, and again he has a lot of the characteristics that future detectives throughout the history of this genre will have. He's eccentric. He has a hobby that he's obsessive about - in this... in his case, it's the love of roses. He's a fanatic about the breeding of roses; and here think of Nero Wolfe and his orchids, Sherlock Holmes and his violin, a lot of those later classic detective heroes have this kind of outside interest that they... they go to as a kind of antidote to the evil and misery they encounter in their daily lives. At one point, Cuff says he likes his roses

because they offer solace, uh, an escape, from the world of crime he typically operates in.

לאורך הטקסט אפשר למצוא פאזות, פילרים (uh, um, well), חזרות (and... and), תיקונים (detective, uh, superintendent Seegrave; which you'll see repeated - um, you'll see this ) (pattern repeated), דוגמאות למשלב דיבורי ( you have the regular cops; they actually make a mess of the investigation). היחידות הרעיוניות קצרות יחסית, ומשורשרות באמצעים פשוטים (and, so). גם ברמה הגלובלית הקטע אינו מאורגן: הדובר פותח בתיאור הגיבור, קף, אך נמלך בדעתו מיד ומחליט לגשת לנושא זה מזווית אחרת; לאחר שהודיע שבדעתו לתאר את כוחות המשטרה בספר, ולפני שעשה זאת בפועל, הוא חוזר אל הנושא הכללי של ההרצאה, השפעתו של "אבן הירח" על ספרות הבלש; הוא נוטש באמצעו את תיאור השפעתו של קף על ארכיטיפ הבלש הספרותי, ועובר להרחיב על אהבתו לוורדים. אמנם ההגייה המוקפדת והמשחק הבינוני גורמים למאפיינים האורליים האלה להישמע מתוסרטים, אבל הם מתוסרטים היטב, ותורמים רבות לאותנטיות של הטקסט.

במהלך ההאזנה מוצגת לנבחנים תמונת אילוסטרציה (אחת או יותר). התמונות אינן גנריות לגמרי, אלא תומכות בטקסט. לדוגמה, במהלך השיחה בין מאמן הכדורסל ובין הסטודנטית מוצגת לנבחנים תמונתם כשהם ישובים במשרדו: המאמן חבוש כובע מצחייה ועונד משרוקית, ועל מדף מונחים כדור וכמה גביעים. פרטים אלו קושרים בין התמונה לטקסט ומסייעים לנבחנים להבין את הקשר השיחה ואת תפקידי הדוברים. במהלך ההרצאה על "אבן הירח" מוצגת תמונה גנרית של מרצה וסטודנטים, אך פעמיים – כאשר המרצה מציג לראשונה את הספר ואת הגיבור – היא מוחלפת לזמן קצר באיור של לוח שנכתבה עליו בגיר הכותרת המתאימה.

### **International English Language Testing System (IELTS)**

מבחן IELTS, הנמצא בבעלות משותפת של המועצה הבריטית, IDP אוסטרליה ומרכז הבחינות ESOL של אוניברסיטת קיימברידג', בוחן שליטה בשפה האנגלית, להלכה בכל הרמות, ומועבר בשתי גרסאות: גרסה אקדמית, המיועדת בעיקר למבקשים ללמוד לימודים גבוהים, וגרסה כללית, המיועדת בעיקר למבקשים להגר למדינה דוברת אנגלית. הוא מורכב מארבעה חלקים –

הבנת הנשמע, הבנת הנקרא, הבעה בכתב והבעה בעלפה – אך החלקים הראשון והאחרון משותפים לשתי הגרסאות. לפרק הבנת הנשמע מוקצות שלושים דקות, ואחריהן עוד עשר דקות להעתקת התשובות אל גיליון התשובות, ויש בו ארבעה טקסטים קצרים: דושיח יום-יומי (לדוגמה, שיחה באשנב המודיעין של הרכבת), מונולוג יום-יומי (לדוגמה, חזרה של מציל על כללי הבטיחות בברכת שחייה), שיחה בהקשר חינוכי כלשהו בין שניים עד ארבעה אנשים (לדוגמה, שיחה בין עוזר הוראה לשני סטודנטים על עבודת-גמר), ומונולוג על נושא אקדמי (לדוגמה, רפרט על נדידת פרפרים). אחדים מהטקסטים מפוצלים לשני חלקים. הנבחנים קוראים את השאלות הרלוונטיות לכל קטע לפני השמעתו, ועונים עליהן במהלכו; כל קטע מושמע פעם אחת בלבד. בסך הכול יש במבחן ארבעים שאלות, עשר לכל טקסט.

ברובם המכריע של הפריטים במבחנים שנסקרו, הנבחנים נדרשים להשלים במילים בודדות (בשום מקרה לא יותר משלוש) פרטי מידע חסרים במערך כלשהו – טבלה, פתק, מחברת וכו'. שאלות אלו – וכמעט כל הפריטים בבחינה – הן שאלות איתור פשוטות: התשובה ניתנת בטקסט במפורש, בניסוח זהה לניסוח המבוקש. יתרה מזו, שאר מערך הנתונים מלא, והדבר מקל על הנבחנים לעקוב אחרי הטקסט. לדוגמה, הקטע הבא מופיע בטקסט השלישי מאחד המבחנים שנסקרו:

Ian: Won't it be hard to write a technical analysis? After all, it's a kids' book.

Tutor: Well, it was initially pitched at adults you know, but over the years it has become seen as a more youth-orientated work. And you're right in a sense – the simple vocabulary and absence of foreshadowing make the story very easy to follow and ideally suited for children. But that doesn't mean there isn't much to analyse. Look at the symbolism, for instance.

Lorna: Symbols are things, right? Material things – like objects – that stand for abstract ideas.

Tutor: Absolutely, yes. And the author uses many of them. There's the robin redbreast, for example, which symbolises the wise and gentle nature that Mary will soon adopt – note that the robin is described as “not at all like the birds in India”. Roses are used as well – as a personal symbol for Mistress Craven – you'll see they're always mentioned alongside her name. And Mistress Craven's portrait can also be interpreted as a symbol of her spirit.



Ian: Are symbols just another name for motifs?

Tutor: No, motifs are a bit different. They don't have as direct a connection with something the way that a symbol does. Motifs are simply recurring elements of the story that support the mood.

Lorna: Are there any in this novel?

Tutor: Yes, two very important ones. The Garden of Eden is a motif. It comes up a few times in connection with the garden of the story. And then you've got the role that secrets play in the story. In the beginning, everything is steeped in secrecy, and slowly the characters share their secrets and in the process move from darkness to lightness, metaphorically, but also in the case of Colin, quite literally. His room in the beginning has the curtains drawn, and he appears at the end in the brightness of the garden.

והוא מתאים לקטע הבא בחוברת הבחינה :

Audience:

Good for children – story simple to follow

Symbols (physical items that represent **25** .....):

- the robin redbreast
- **26** .....
- the portrait of Mistress Craven

Motifs (patterns in the story):

- the Garden of Eden
- secrecy – metaphorical and literal transition from **27** .....

בדוגמה לעיל, הנבחנים חייבים לעקוב אחרי הטקסט כולו, ואינם יכולים להשתמש במילים שבפריט עצמו כעין דגלים שיסמנו את התשובה הנכונה – מפני שהמשפטים בפריט הם פרפרזה של המשפטים בטקסט (למשל, "physical items that represent" במקום "material things - like " objects - that stand for", ומפני שלעתים המילים בפריט מופיעות בטקסט אחרי התשובה (למשל, המילים "metaphorically, but also... quite literally" מופיעות בטקסט אחרי המילים "darkness to lightness"). המערך המאורגן והמלא-למחצה מספק לנבחנים חלשים עוגנים

להיאחז בהם אם איבדו את הדובר. עם זאת, הוא מקל מאוד על נבחנים חזקים יותר – ועלול לעודד התרחקות של הטקסטים עצמם מהאותנטיות הרצויה. טקסטים דבורים אמתיים בדרך כלל אינם בנויים לתלפיות, ולעתים רחוקות בלבד אפשר לסכם את הנאמר בהם בטבלה. ההגייה הנקייה והיעדר ההיסוסים והשגיאות מחריפים מגמה זו.

בשאלות הפתוחות, הנבחנים נדרשים להקפיד על איות ודקדוק תקינים – גם כאשר שגיאות אינן מעידות על הבנה לקויה של הנאמר.

### **Pearson Test of English General (PTEG)**

מבחן PTEG בוחן שליטה בשפה האנגלית בשימוש כללי, ומועבר בשש רמות, המקבילות לשש דרגות השליטה בשפה לפי המדד האירופי להערכת שפות (CEFR). הוא מורכב משלושה עשר חלקים: מרביתם מוקדשים לאחת מארבע מיומנויות השפה, ומקצתם משלבים ביניהן. הבנת הנשמע נבחנת בפרקים 1–3, ואורכם הכולל 15–25 דקות, לפי רמת הבחינה.

בפרק הראשון, הנבחנים שומעים עשרה טקסטים קצרים ממגוון סוגים: דיאלוגים, עדויות, תשדירים, הודעות, פסקאות אינפורמטיביות ועוד. כל טקסט מושמע פעם אחת בלבד, ואחריה על הנבחנים להשיב על שאלה אחת ולה שלוש אפשרויות תשובה. לפני השמעת הטקסט מוקצבות עשר שניות לקריאת השאלה. מטרת הפרק לבחון את יכולתם של הנבחנים להבין את הרעיון המרכזי ואת הפרטים העיקריים בטקסטים דבורים קצרים, וחלק מהשאלות משלבות גם היסקים פשוטים, למשל על הלך הרוח של הדוברים. מיעוט קטן מהשאלות דורשות תשומת-לב לפרטים ולשמות, אך מרביתן אכן דורשות הבנה גלובלית, או לכל הפחות קליטת פרטים שאפשר לצפות מהמאזין לקלוט גם בהאזנה שאיננה מכוונת-מטרה – כך שאפשר לחלוק על הבחירה להקצות זמן לקריאת השאלות מראש. לדוגמה, שתי השאלות להלן, מתוך מבחן ברמה B2, אינן דורשות בהכרח הצגה מראש של מטרת-האזנה:

(הנבחנים קוראים):

The main attraction of the Oxford Fashion Museum's new exhibition is

- (A) footwear.
- (B) headwear.
- (C) swimwear.

(הנבחנים שומעים):

**Listen to the speaker. What is the main attraction of the Oxford Fashion Museum's new exhibition?**

M: The Oxford Fashion Museum's 2010 exhibition opens on January 10th. There are sections on shoes, boots and other footwear borrowed from museums across Europe and, of course, the long-awaited collection of Victorian hats which will attract thousands. But the biggest draw for visitors has to be the room full of the bathing costumes people have worn on the beach from the last two centuries.

(הנבחנים קוראים):

What are the speakers doing?

- (A) complaining
- (B) discussing
- (C) disagreeing

(הנבחנים שומעים):

**Listen to the conversation. What are the speakers doing?**

F: I can't believe some of the things people wear today! It's so different from when we were growing up.

M: I know. I was out in town last night. It's the middle of winter and there were girls wearing short skirts and T-shirts.

F: And no coats either, I imagine. It seems to be the fashion these days...

הפרק השני הוא הכתבה. הנבחנים שומעים טקסט קצר פעמיים: בפעם הראשונה הוא מושמע ברצף, ובשנייה – במקוטע, עם הפסקה בת 15 שניות אחרי כל מקטע, שבה על הנבחנים להעלות על הכתב את שנאמר. הנבחנים נדרשים לכתוב את הטקסט כלשונו, והדבר מחייב אותם לזכור לא רק את תוכן הדברים, אלא גם את הניסוח שלהם; בחירה זו פוגעת באוטנטיות של המטלה (שכן במשימות האזנה אמתיות בדרך כלל אין חשיבות לזכירת הניסוח המדויק), ומגבילה מאוד את אורך המקטעים, ולפיכך גם את מידת ההסתמכות של המטלה על הבנה ברמה גבוהה. להלן טקסט לדוגמה מתוך בחינה ברמה C2, ובו ההפסקות מסומנות בלכסן:

Some claims concerning the merits of ecotourism / are exaggerated, / or owe more to labelling and marketing / than genuine sustainability. / Not only are such projects frequently planned / and implemented without local consent, / they often threaten local cultures / and natural resource bases. / Critics regard this niche tourism / as a façade, / a tactic concealing / the mainstream tourism industry's exploitative practices.

אף שמדובר בבחינה ברמת הקושי הגבוהה ביותר, הרוב המכריע של המקטעים אינם תובעניים כלל מבחינת העומס הקוגניטיבי שהם יוצרים. באחדים מהם, למשל "are exaggerated" או "as

"a façade", נראה שהקושי המרכזי הוא האיות (הנבחנים נקנסים על שגיאות איות, בהתאם לתפישת השילוב בין מיומנויות). נראה שהמגבלה על אורך המקטעים היא גם הגורם לכך שחלק

מההפסקות משולבות במקומות לא טבעיים, למשל באמצע "planned and implemented".

קוצר המקטעים והעובדה שבחירתם אינה עולה תמיד בקנה אחד עם החלוקה התחבירית של המשפט מדגישים את הסתמכותה של המטלה על כישורים מכניים לא פחות מאשר על הבנה.

הפרק השלישי מכיל שני טקסטים בינוניים באורכם – ריאיון קצר, הרצאה קצרה וכו' – שלכל אחד מהם מצורפות חמש שאלות איתור בפורמט השלמה. כל טקסט מושמע פעמיים רצופות,

ולנבחנים ניתנת הזדמנות לקרוא מראש את השאלות הרלוונטיות. כמו במבחן IELTS, גם כאן

המטלה אינה דורשת היסקים או עיבוד גלובלי – החלקים החסרים ניתנים בטקסט במפורש

ובניסוח המבוקש – אך ניסוח השאלות מחייב את הנבחנים לעקוב אחרי הטקסט. לדוגמה,

Responsible travellers prefer learning and discovering rather than "

If you're after fulfilment, adventure " completing a \_\_\_\_\_

and knowledge - instead of simply ticking off a travel to-do list - then responsible

"travel is for you

הטקסטים בבחינה מגוונים, משקפים שלל סיטואציות שימוש אמיתיות, ומוגשים במגוון מבטאים.

עם זאת, הם עניים במאפיינים של שפה דבורה.

### **The European Language Certificates (TELC)**

מבחני TELC מועברים באחת-עשרה שפות ובעשרות גרסאות. במרבית השפות נבחנת רק שליטה

בשפה בשימוש כללי, והגרסאות נבדלות זו מזו רק ברמת השליטה הנבחנת (לדוגמה, באיטלקית

וברוסית יש ארבע בחינות, המתאימות לרמות B2–A1 לפי המדד האירופי להערכת שפות; בצ'כית, בערבית ובפורטוגזית יש רק בחינה אחת, ברמה (B1). בטורקית, בצרפתית ובספרדית יש גם גרסאות מותאמות לגיל בית הספר. בכמה שפות, ובראש ובראשונה באנגלית ובגרמנית, יש גם גרסאות מותאמות לנושאים ספציפיים, כמו לימודים אקדמיים, עסקים, רפואה, סיעוד ותיירות. במרבית גרסאות הבחינה חמישה פרקים, המוקדשים לכל אחת מארבע מיומנויות השפה ולאבני בניין (אוצר מילים ודקדוק). לפרק הבנת הנשמע מוקדשות 15–60 דקות, לפי רמת הבחינה. גם מבנה הפרק משתנה בין הרמות השונות.

ברמה C2, פרק הבנת הנשמע כולל מטלה אחת בלבד: סיכום הרצאה בת כרבע שעה. לכתובת הסיכום מוקצות 45 דקות, והנבחנים נדרשים להציג בו את כל חמשת הרעיונות המרכזיים שעלו בהרצאה, ולפחות שתי תת-נקודות לכל רעיון מרכזי. מילוי דרישה זו יקנה לנבחן עד 25 נקודות. נוסף על כך, הנבחנים מקבלים ציון גם על הדרך שהביעו בה את תוכן ההרצאה (0, 5, 10 או 15 נקודות); ציון זה הוא גלובלי, וניתן לפי בהירות הטקסט, מידת הקוהרנטיות והדיוק שהוצג בהם הציר הרעיוני של ההרצאה, וכן האיות, התחביר, הדקדוק ואוצר המילים. רמת ההרצאות גבוהה יחסית, הן מבחינת התוכן הן מבחינת הלשון, והן קרובות לקצה ה"כתוב" של סקלת הדיבוריות. הדבר בא לידי ביטוי גם במידה גבוהה של ארגון רעיוני, המאפשרת לשלב בהרצאה התייחסויות ישירות רבות למבנה שלה, ובכלל זה התוויה מפורשת שלו בתחילת ההרצאה. שלטי ההכוונה המטפוריים האלה מקלים על הנבחנים לעקוב אחרי הנאמר, ואף מאפשרים לנבחנים שאיבדו את הקו הרעיוני של המרצה לשוב אליו אחר כך. לדוגמה, פסקה ראשונה מתוך הרצאה (ההדגשות שלי; יג"מ):

Thank you, Professor White. It's a great pleasure for me to be here today. As you know, I'll be discussing artificial intelligence, or AI. There's a lot of ground to cover and the research is extensive, so *I'll be focussing on a few key questions. To begin with*, I will speak about the activity many thinkers have claimed is distinctively human, namely the ability to speak and communicate using language. To what extent are computers able to process and generate human language? *I'll then turn to* ludic behaviour, something we humans share with much of the animal world - that is,

engaging in games and play. Computers may be able to outperform humans on certain tasks, but to what extent are machines able to learn while engaged in play as we do? Anyone who has observed a child at play will know that humans can enjoy almost limitless creativity. *In the third part of my talk* I will address the extent to which machines can be considered "creative" or to have the power of imagination. *From there I will move on to* discussing what role emotions might play in artificial intelligence *and, finally, end by* tackling the question of what might stand in the way of assembling artificial counterparts of ourselves.

כאמור, גם המשך ההרצאה עשיר בסמני-מבנה, החל בחלוקה ברורה לפסקאות (באמצעות הפסקות ארוכות יחסית בדיבור), דרך שילוב כותרות-משנה בטקסט (למשל באמצעות שאלות שהתשובה עליהן ניתנת מיד) וכלה בהתייחסויות מפורשות לחלוקה הנושאית ( "and this leads me to the last point I wanted to mention with regard to artificial speech", "this brings me to my third point: can computers ever become truly creative?").

גם ברמה C1 יש מטלת סיכום הרצאה, אלא שברמה זו פרק הבנת הנשמע כולל עוד שתי מטלות. המטלה הראשונה מורכבת משני חלקים. בחלק הראשון הנבחנים שומעים שמונה מונולוגים קצרים (כ-45 שניות) באותו נושא, ונדרשים לקבוע לגבי כל אחד מה עמדת הדובר באותו נושא, מבין שלוש אפשרויות קבועות (לדוגמה, האם הוא מעדיף לחזור לגור במולדתו, להישאר בנכר, או שאין לו העדפה מיוחדת לכאן או לכאן). בחלק השני הנבחנים קוראים עשרה משפטים קצרים, מאזינים שוב לשמונת המונולוגים, ומתאימים משפט אחד לכל מונולוג. להשלמת המטלה נדרשת הבנה גלובלית של הנאמר: עמדת הדובר והרעיון המרכזי בדבריו.

במטלה השנייה בפרק, הנבחנים מאזינים לריאיון רדיופוני (כ-4.5 דקות) ומשלימים בשעת ההאזנה עשרה משפטים שקראו מראש (לא יותר מארבע מילים לכל השלמה).

המטלה השלישית בפרק היא, כאמור, מטלת סיכום הרצאה. ההרצאה קצרה מההרצאה ברמה C2 (כעשר דקות), והנבחנים אינם נדרשים להציג את הסיכום כטקסט שלם וקוהרנטי, אלא להשלים רובריקות נתונות מראש. לדוגמה, הקטע הבא:

So, now you know what it is – where did it all begin?

Well, some people say that the Americans were first with an organisation called “Ten Thousand Villages” which began buying needlework from Puerto Rico in 1946. Ten years later, the first formal “Fair Trade” shop which sold these and other items opened in the USA.

Fair Trade in Europe dates from the late 1950’s when the charitable organisation Oxfam UK started to sell crafts made by Chinese refugees in Oxfam shops. In 1964 it created the first official Fair Trade Organization.

At the same time, Dutch third world groups began to sell cane sugar with the message “by buying cane sugar you give people in poor countries a place in the sun of prosperity”. These groups went on to sell handicrafts from the developing world and in 1969 the first “World Shop” opened. There are now many such shops throughout Europe. In 1973, the first “fairly traded” coffee was imported to Europe from cooperatives of small farmers in Guatemala. Over the next thirty years this gradually became a concept with hundreds of thousands of coffee farmers benefiting from it. Now Fair Trade coffee makes up between 25 and 50% of the turnover of Fair Trade Organisations in Europe today. Which brings me back to the Fair Trade Mark I mentioned earlier. In the mid-1980s, a priest working with smallholder coffee farmers in Mexico had the idea of using some kind of system to identify Fair Trade products. Products bought, traded and sold fairly would qualify for a label that would make them stand out among ordinary products on store shelves. In 1988, the Fair Trade Mark was created. The concept caught on: within a year, coffee with the label had a market share of almost three percent.

מספק את התשובות לחלק הבא בחוברת הבחינה :

**b) The history of the Fair Trade Association:**

1940's: \_\_\_\_\_

1950's: \_\_\_\_\_

1960's: \_\_\_\_\_

1970's: \_\_\_\_\_

1980's: \_\_\_\_\_

המטלות השנייה והשלישית בפרק מזכירות במידה רבה את המטלות במבחן IELTS שנסקר לעיל: משימות איתור לאטריוויאליות, הדורשות מהנבחן לעקוב אחר הטקסט כולו ולהבינו.

בפרקי הבנת הנשמע בבחינות ברמות הביניים (B2–B1) שלושה חלקים, ובהם רק שאלות נכון/לא נכון. בחלק הראשון חמש שאלות הבוחנות את הבנת הלזו של טקסטים קצרים; בחלק השני עשר שאלות הבוחנות הבנה לוקלית ומתייחסות לטקסט ארוך יחסית; ובחלק השלישי חמש שאלות הבוחנות שלילת פרטים ספציפיים מתוך טקסטים קצרים. מלבד הרמה הלשונית הכללית, הבחינות נבדלות זו מזו ברמת הדיבוריות של הטקסטים: לדוגמה, ידיעות מתוך מבזק חדשות לעומת מונולוגים של אנשים מן היישוב בחלק הראשון, וריאיון רדיופוני לעומת שיחת חולין בחלק השני. כמורכן, ברמה B1 הטקסטים בחלקים השני והשלישי מושמעים פעמיים, וברמה B2 – רק פעם אחת. הדמיון המבני בין שתי הבחינות מעלה חשד שלהשמעה החוזרת בבחינה ברמה B1 אין הצדקה פסיכומטרית של ממש מלבד הקלה על הנבחנים.

ברמות הנמוכות אין אחדות מבנית בין הגרסאות בשפות הבחינה השונות. יצוין כי בחלק מהגרסאות יש חלק המוקדש לשאלות דיסקרטיות (לדוגמה, הנבחנים שומעים " How much did your holiday flat cost? - It was £315 a week", ונדרשים להכריע אם המחיר הוא 315 ליש"ט, או 350). כאמור, פריטים מסוג זה בדרך כלל אינם אותנטיים, שכן עיבוד השפה נעשה "מלמעלה למטה" לא פחות מאשר "מלמטה למעלה". ברם בדוגמה שהובאה, וכן בשאר השאלות בבחינה, אין למאזין בררה אלא להסתמך על העיבוד הפונולוגי – אין רמזים סמנטיים או תחביריים, למשל, שיסייעו לו להכריע בין 315 ובין 350. יש אפוא בשאלות אלו כדי לבחון את שליטתם של



נבחנים חלשים באלמנטים הבסיסיים ביותר של השפה.

הבחינות המותאמות לנושאים ספציפיים שונות מבחינות השימוש הכללי בשפה בראש ובראשונה בעולם התוכן של הטקסטים המובאים בהן. עם זאת, העובדה שלבחינה ייעוד ספציפי מאפשרת דיוק רב יותר גם בבחירת המטלות וסוגי הטקסטים, והתאמתם לסיטואציות ההאזנה הספציפיות שאנו מצפים שהנבחנים ייתקלו בהן. לדוגמה, בבחינת הגרמנית לעובדי סיעוד, הנבחנים נדרשים להתמודד עם הודעות טלפוניות, שיחות ושיבת צוות המכילות דיווחים על מצבם של מטופלים, הוראות לטיפול עתידי בהם וכיוצא בזה.

בכל גרסאות הבחינה שנסקרו הגיית הטקסטים מוקפדת. מלבד שימוש ניכר ומעט מודגש באינטונציה, כמעט ואין מאפיינים אורליים מובהקים.

### **Test of English for International Communication (TOEIC)**

מבחן TOEIC נועד להעריך את היכולת לתקשר בשפה האנגלית בחיי היומיום בסביבה בינלאומית, בעיקר בעולם העסקים, המסחר והתעשייה. לפרק הבנת הנשמע בו מוקצות 45 דקות, ויש בו מאה שאלות, המחולקות לארבעה חלקים: עשר שאלות התאמת תיאור לתמונה, שלושים הערכות תגובה, שלושים שאלות על עשרה דיאלוגים קצרים, ושלושים על עשרה מונולוגים קצרים.

בחלק הראשון, הנבחנים מתבוננים בתמונות בחוברת הבחינה, ושומעים ארבע הצעות לתיאור כל תמונה. עליהם לבחור בתיאור המתאים לתמונה. התשובה הנכונה איננה בהכרח סיכום קולע של הנראה בתמונה; לדוגמה:



- (A) He's checking his watch.
- (B) He's wearing a jacket.
- (C) He's adjusting his tie.
- (D) He's folding his clothes.

בחלק השני, הנבחנים שומעים שאלה או אמירה, ואחריה שלוש תגובות; עליהם לבחור בתגובה ההגיונית. לדוגמה:

They've reduced the price of these cameras.

(A) Then let's go ahead and buy one.

(B) I don't think we've been introduced.

(C) I never win anything.

בשני סוגי המטלות, לכאורה נבחנת הבנה ברמת היחידה הרעיונית; בסוג השני יש רמז גם לכישורים קומוניקטיביים. ואולם, המטלות קרובות במידה מסוכנת לדיסקרטיות: לדוגמה, מסיח ג במטלה השנייה נשען על ההבדל בין price ובין prize. אפשר לבנות בפורמטים האלה גם מטלות מאתגרות ואותנטיות יותר.

בחלקים השלישי והרביעי, הנבחנים שומעים דיאלוגים ומונולוגים קצרים, ועונים על שלוש שאלות בררה הנוגעות לנקודות מפתח בהם. השאלות נתונות בחוברת הבחינה, אך לא ניתן לנבחנים זמן לקרוא אותן לפני ההשמעה; מוטב כד, שכן השאלות נוגעות ללוח של הנאמר, והנבחנים אמורים לדעת להשיב עליהן גם בלי שידעו מראש מה לחפש. לדוגמה, לדיאלוג הבא:

(Woman) I think I'll have to take the train to the regional sales meeting up in the city next week.

(Man) Don't you usually drive when you go to those meetings? I thought you didn't like to take the train.

(Woman) I don't, but the highway's being repaired, and I'm afraid I might be late if I have to make a detour through an area I don't know very well.

(Man) You're right. And it'll be expensive to park up there, too.

מוצמדות השאלות "Why is the woman going to the city", "How will she get there", ו"What is the problem". השאלות בחלקים אלו אותנטיות יותר מבשני החלקים הראשונים, ובוחנות הבנה ברמה גבוהה יותר: בין היתר נדרשת הבנה של רעיונות מרכזיים, השתמעויות וכוונות. קוצר הפריטים מונע בנייה של קונטקסט רחב או של פריטים הדורשים הבנה מעמיקה וסינתזה, אך כנגד זה מאפשר שילוב מספר גבוה של פריטים בבחינה, וגיוון תמטי ניכר ביניהם; וממילא ספק אם הבנה מעמיקה מסוג זה היא נדבך מרכזי במבנה שהמבחן מבקש להעריך.

## Test Deutsch als Fremdsprache (TestDaF)

מבחן TestDaF בוחן שליטה בשפה הגרמנית ברמה גבוהה (מקביל לרמות C1–B2 לפי המדד האירופי להערכת שפות), ומיועד בראש ובראשונה למעוניינים ללמוד במדינות דוברות גרמנית. הוא מורכב מארבעה חלקים: הבנת הנקרא, הבנת הנשמע, הבעה בכתב והבעה בעלפה. לפרק הבנת הנשמע מוקצות שלושים דקות, ואחריהן עוד עשר דקות להעתקת התשובות אל גיליון התשובות, ויש בו שלושה טקסטים ברמות שונות המתאימים תמטית ולשונית לעולם הלימודים הגבוהים: דיאלוג אופייני לחיי היומיום באוניברסיטה, ריאיון רדיופוני עם שלושה או ארבעה דוברים, והרצאה קצרה או ריאיון עם מומחה. לפני כל טקסט מושמע תיאור קצר של הסיטואציה, המופיע גם בכתב בחוברת הבחינה, ואחריו הנבחים מקבלים זמן לקרוא את השאלות הנלוות לטקסט.

הטקסט הראשון נועד לבחון את יכולתם של הנבחים להתמודד עם סיטואציות תקשורתיות מחיי היומיום, למשל שיחה בין פרופסור לתלמיד על עתידו האקדמי. על הנבחים לענות בקצרה על שמונה שאלות איתור פתוחות ("מדוע הסטודנט רוצה לגשת לבחינה מוקדם מהרגיל?", "איזה קורס שפה חסר עדיין לסטודנט?" וכו'). הטקסט (כשתי דקות) מושמע פעם אחת, והנבחים עונים על השאלות במהלכו.

הטקסט השני בוחן את היכולת לחלץ מידע חשוב מהטקסט ולעבד אותו תוך לקיחה בחשבון של מבעים קודמים בטקסט. על פי רוב הוא עוסק בעולם האקדמי באופן כללי, למשל רבישיח על האופק המקצועי של דוקטורנטים, או על דרכי הסינון של מועמדים ללימודים אקדמיים. על הנבחים לענות על עשר שאלות נכון/לא-נכון. שאלות רבות עוסקות בעמדותיהם של המידיינים, והתשובות להן אינן מובאות בטקסט ישירות, אלא משתמעות ממנו. לדוגמה, את התשובה לשאלה "ד"ר קלנק מאמין שקריירה לא צריכה להיות המניע היחיד להשגת דוקטורט" הנבחים נדרשים להסיק מהציטוט הבא מפיו: "[...] לי עצמי היו תקופות שרציתי לפרוש. אבל אני נהנה מעבודה מדעית, לכן תמיד המשכתי. מי שאין לו מניע כזה, וכותב דוקטורט רק למען הקריירה, עלול לנשור בקלות". הטקסט (כארבע דקות) מושמע פעם אחת, והנבחים עונים על השאלות במהלכו.

הטקסט השלישי בוחן את היכולת לעקוב אחרי טקסט מורכב בנושא מדעי ולהשיב בקצרה על שאלות הנוגעות לעניינים מרכזיים בו. כאמור, הנושא מוצג בהרצאה קצרה או בריאיון עם מומחה. על הנבחים לענות בקצרה על שבע שאלות פתוחות. השאלות קשות יותר מהשאלות

המלוות את הטקסט הראשון, ודורשות אינטגרציה של מידע והבנה גלובלית של מקטעים ארוכים יחסית בטקסט. הטקסט (כחמש דקות) מושמע פעמיים, עם הפסקה בת דקה בין שתי ההשמעות. הנבחים מונחים לענות על השאלות כבר במהלך ההאזנה הראשונה.

הטקסטים מגוונים ונראה שהם מייצגים היטב את סיטואציות ההאזנה האופייניות ללימודים אקדמיים. כולם ארוכים יחסית, והדבר מאפשר להם לבנות קונטקסט אמין וליצור סימולציה טובה של שימוש מכוון-מטרה בשפה. עם זאת, ניכר שהם מתוסרטים: ההגייה בהם מוקפדת ואין בהם היסוסים או תיקונים.

פריטי המבחן והתשובות שהנבחים צריכים להפיק כולם בגרמנית, כמתחייב במבחן בין-לאומי גדול. עם זאת, נעשה מאמץ למזער את מידת ההתערבות של משתנים לא-רלוונטיים: השאלות כולן קצרות ומורכבות ממשפט אחד, ללא אפשרויות תשובה, שקריאתן גוזלת זמן ומאמץ; וכאשר הנבחים מתבקשים לייצר בעצמם תשובה, הם מונחים לכתוב בקיצור נמרץ.

- Blau, E. (1990). "The Effect of Syntax, Speed and Pauses on Listening Comprehension". *TESOL Quarterly*, 24, 746-753.
- Bozorgian, H. (2012). "The Relationship Between Listening and Other Language Skills in International English Language Testing System". *Theory and Practice in Language Studies*, 2(4), 657-663.
- Brindley, G. & Slatyer, H. (2002). "Exploring Task Difficulty in ESL Listening Assessment". *Language Testing*, 19(4), 369-394.
- Buck, G. (2001). "Assessing Listening". Cambridge: Cambridge University Press.
- Duker, S. (1965). "Listening and Reading". *The Elementary School Journal*, 65(6), 321-329.
- Field, J. (2011). "Into the Mind of the Academic Listener". *Journal of English for Academic Purposes*, 10, 102-112.
- Filipi, A. (2012). "Do Questions Written in the Target Language Make Foreign Language Listening Comprehension Tests More Difficult?". *Language Testing*, 29(4), 511-532.
- Ginther, A. (2001). "Effects of the Presence and Absence of Visuals on Performance on TOEFL CBT Listening-Comprehensive Stimuli" (TOEFL Research Report No. 66).
- Liu, J. & Costanzo, K. (2013). "The Relationship Among TOEIC Listening, Reading, Speaking and Writing Skills". In Powers, D. (Ed.), *The Research Foundation for the TOEIC Tests: A Compendium of Studies (Vol. II)*, 2.1-2.25.
- Lynch, T. (2011). "Academic Listening in the 21st Century: Reviewing a Decade of Research". *Journal of English for Academic Purposes*, 10, 79-88.
- Powers, D. (1986). "Academic Demands Related to Listening Skills". *Language Testing*, 3(1), 1-38.
- Powers, D. (2013). "Assessing English-Language Proficiency in All Four Language Domains: Is It Really Necessary?". In Powers, D. (Ed.), *The Research Foundation for the TOEIC Tests: A Compendium of Studies (Vol. II)*, 1.1-1.7.

- Powers, D., Yu, F. & Yan, F. (2013). "The TOEIC Listening, Reading, Speaking, and Writing Tests: Evaluating Their Unique Contribution to Assessing English-Language Proficiency". In Powers, D. (Ed.), *The Research Foundation for the TOEIC Tests: A Compendium of Studies (Vol. II)*, 3.1-3.13.
- Richards, J. C. (1983). "Listening Comprehension: Approach, Design, Procedure". *TESOL Quarterly*, 17(2), 219-240.
- Rosenfeld, M., Leung, S. & Oltman, P. (2001). "The Reading, Writing, Speaking, and Listening Tasks Important for Academic Success at the Undergraduate and Graduate Levels" (TOEFL Monograph Series Report No. 21).
- Sawaki, Y., Stricker, L. & Oranje, A. (2009). "Factor Structure of the TOEFL Internet-Based Test". *Language Testing*, 26(1), 5-30.
- Shohamy, E. & Inbar, O. (1991). "Validation of Listening Comprehension Tests: The Effect of Text and Question Type". *Language Testing*, 8(1), 23-40.
- Taylor, L. & Geranpayeh, A. (2011). "Assessing Listening for Academic Purposes: Defining and Operationalising the Test Construct". *Journal of English for Academic Purposes*, 10, 89-101.
- Wagner, E. (2007). "Are They Watching? Test-Taker Viewing Behavior During an L2 Video Listening Test". *Language Learning & Technology*, 11(1), 67-86.
- Wagner, E. (2010). "The Effect of the Use of Video Texts on ESL Listening Test-Taker Performance". *Language Testing*, 27(4), 493-513.