

## אמל"ק

מפאת מנגישה לציבור מאגר מתויג (dataset) של כ-300,000 משפטים ב-NLI. המאגר מבוסס על multilinli, שעבר תרגום אוטומטי ע"י מודל שפה גדול - Gemini ובחינה לוגית ולשונית מעמיקה. בעבודה ויידאנו שהתרגום לא משפיע/משנה את התיוג המקורי, ושהמשפט המתורגם נשמר בשפה טבעית. כמו כן בחנו מהי הדרך האופטימלית כיום לתרגם משפטים. היות ש-multilinli מורכב ממספר מקורות מידע (ג'אנרים) שונים, בוצעו תרגום ובדיקות אימות נפרדות לכל ז'אנר, ונפו ז'אנרים שתרגומם לא היה טוב מספיק. המאגר המתורגם, התרגומים לכלל מאגר הבסיס - multilinli, וה-prompts שבהם השתמשנו לתרגום, כמו גם תיעוד התהליך המפורט בפרסום זה זמינים ב-github שלנו <https://github.com/NNLP-IL/HebNLI>, וב-hugging face, <https://huggingface.co/datasets/HebArabNlpProject/HebNLI>.

## רקע

פרויקט היסק שפה טבעית בעברית הוא חלק מתוכנית ה-NLP הלאומית בעברית וערבית. במסגרת זו, מובילה מפא"ת תוכנית הכוללת מספר רב של פרויקטים, ע"מ לפתח אבני יסוד לעיבוד שפות טבעיות, ולהעמיד אותן לרשות קהילת העוסקים בתחום. אבני היסוד כוללות בניית מאגרי מידע מתויגים, מודל שפה ענק (LLM), ופיתוח יכולות ושירותים מתקדמים בעברית וערבית בקוד פתוח (open source) כשירות למדענים, לארגונים ממשלתיים ולחברות מסחריות. כלל התוצרים המפותחים פתוחים לציבור ללא עלות.

פרויקט ה-NLI מנגיש לציבור מאגר מתויג (dataset) מגוון ורחב היקף של משפטים טבעיים ע"ב עברית ותיוגם ברמת ביצוע גבוהה. לביצוע הפרוייקט מפא"ת פנתה לחברת Webiks.

חברת Webiks מתמחה בביצוע פרויקטים בגישת Data Centric AI ובפרט בהכנת דאטה-סטטים טקסטואליים מתויגים באיכות גבוהה בעברית ובערבית, הדורשים יכולות ומתודולוגיות תיוג ובקרת איכות, יכולות ניתוח בלשניות לצד מימוש ושימוש במודלי שפה לחילוץ, ניתוח ותרגום מידע - יכולות נדרשות לפרוייקט ה-NLI.

בפרוייקט יונגשו 303,383 זוגות משפטים בשפה טבעית, מתוך Multilinli Dataset, מתורגמים לעברית ומתויגים, לאחר בחינה וניפוי.

## מהו NLI ומדוע הוא חשוב

משימת ההיסק שפה טבעית (NLI - Natural Language Inference) היא משימה בסיסית בהבנת שפה טבעית על ידי מחשב. משימת ה-NLI בודקת את היכולת של המודל להבין את המשמעות הסמנטית של משפטים בשפה טבעית על ידי בחינת היחס הלוגי שביניהם. זוהי יכולת חשובה ומרכזית בתהליכי הבנת שפה ממוחשבים, ונחשבת ליכולת לשונית גבוהה.

היחס הלוגי בין משפט הנחה (Premise) ומשפט השערה (Hypothesis) יכול להיות יחס גרירה (Entailment) – ניתן להסיק את משפט ההשערה ממשפט ההנחה, יחס סתירה (Contradiction) או יחס ניטרלי (Neutral).  
רוב מאגרי הנתונים ל-NLI מכילים זוגות של משפטים ותיג ע"פ קטגוריות "סותר", "ניטרלי" או "מחזק". ביצועים טובים במשימה זו נחשבים היום לאבן דרך בהבנת שפה טבעית.

### מדוע בחרנו ב-Multinli

Multinli - Multi-Genre Natural Language Inference הוא אחד ממאגרי הנתונים הגדולים והחשובים ביותר עבור מחקר ופיתוח של מודלים להיסק שפה טבעית (NLI) באנגלית.  
MNLI פותח על ידי חוקרים מאוניברסיטת סטנפורד בשנת 2018, והוא מכיל כ-433,000 זוגות משפטים מתורגמים מ-10 ז'אנרים טקסטואליים שונים כמו חדשות, סיפורת, טקסטים אקדמיים ועוד. כל זוג משפטים מורכב ממשפט מקור ומשפט מסקנה, וכולל בין היתר את הצבעות המתייגים, התיג שנבחר (הקשר בין משפט 1 למשפט 2 בכל רשומה - האם המסקנה סותרת, ניטרלית, או מחזקת את המשפט המקורי) ועוד.

**Multinli** משמש לאימון והערכה של מודלי NLI במגוון רחב של נושאים וז'אנרים. זמינותו הרחבה של המאגר תרמה רבות להתקדמות המחקר בתחום ההיסק האוטומטי מטקסטים. העובדה כי המשפטים מגיעים ממגוון רחב של מקורות, סוג שיחה, מין דובר וכו' היוותה יתרון משמעותי.

### תפיסת פתרון:

היות ותיג מספר גדול של משפטים בעברית הינו יקר מאד ומורכב מאד, בחרנו לתרגם dataset זה תוך ביצוע פעולות נוספות להוכחת תקפות הפתרון.

כדי שנוכל להנגיש לציבור מאגר מתיג ברמה גבוהה, נדרשנו לענות על 2 תנאים:  
1. לוודא כי התרגום מגיע לרמה של שפה טבעית בעברית. לצורך זה בחנו מספר אפשרויות תרגום וניסוח, ובדקנו מדגמית את המשפטים שהתקבלו. הבדיקה נעשתה ע"י בלשנים.  
2. להראות כי התרגום לעברית לא שינה את התיג המקורי. כלומר - התרגום לא שינה את משמעות המשפט כך שלדוגמה אחרי תרגום משפט תויג במקור לתיג גרירה לא הפך לניטרלי וכו'.

### שלבי הבדיקה:

**שלב ראשון** - בחינת הנחת היסוד שאמרה שתרגום לא משנה את התיג, ובחירת תרגום מיטבי לרשומות. הבחינה התבצעה על 500 רשומות.

**ניפוי רשומות:** בשלב הראשון נבדק באופן ידני מדגם של כמה מאות רשומות. נמצא כי כ-27% מהרשומות (187 מתוך 690 רשומות) אינן מתאימות לתיג NLI: חלקן אינן קוהרנטיות (כבר בשלב זה היה ברור כי מקור של שיחות מתומללות לא מתאים) וחלקן מורכבות משאלות ולא משפטי חייווי. לאחר הניפוי נמצאו 503 רשומות מתאימות לתיג.

### תרגום אוטומטי ע"י Google Translate

**תיקון התרגום האוטומטי (Google Translate):** בשלב השני נבדק התרגום האוטומטי, ונערכו בו תיקונים. עבור כל רשומה נבדק אם התרגום האוטומטי מובן לקריאה כשלעצמו (ניתן להבינו מבלי להיעזר במקור)

ואם הוא אינו משנה את המשמעות של המשפט באנגלית. תרגום שעומד בשני הקריטריונים האלה מוגדר תרגום עובר סף.

תרגום אוטומטי (דרך Google Translate) של כ-54% מהרשומות נמצא עובר סף.

**תיוג הרשומות המתורגמות:** בשלב השלישי החומר תויג מדגמית על ידי שני מתייגים שונים:

1. מתייג עברית א': 500 רשומות תויגו על פי התרגום המתוקן

2. מתייג עברית ב': 200 רשומות (מתוך ה-500) תויגו על פי התרגום האוטומטי

נערכו השוואות בין שני מתייגי העברית, ובין כל אחד מהם למתייגי האנגלית. התוצאות מוצגות בטבלה הבאה:

סוג השוואה	אחוז הסכמה
השוואה בין מתייגי עברית (א' מול ב')	85.16%
השוואה בין מתייג עברית א' למתייגי המקור (האנגלית)	82.70%
השוואה בין מתייג עברית ב' למתייגי המקור (האנגלית)	80.36%

התובנות משלב זה:

- התרגום לעברית מאנגלית לא משנה את התיוג ב NLI - ניתן לתרגם dataset זה ולשמור על איכות התיוג
- תרגום ע"י google translate בלבד לא מספיק. נדרש שיפור התרגום.

### שלב שני - בחירת מודל תרגום מיטבי:

כאמור, היות שכלי התרגום Google Translate נמצא לא מוצלח דיו עבור מטלת תרגום כל המשפטים הדרושים לפרויקט, הוחלט על בחינתם של כלי תרגום אוטומטיים אחרים. שני כלי התרגום האוטומטיים שנבדקו עבור מטלת התרגום הינם Claude של חברת Anthropic, Bard-I של חברת Google.

התרגום דרך LLM נבחן ב-2 אופנים:

- א. שיפור התרגום שהתקבל מ-Google Translate - הפרומפט הכיל את המשפט העברי שסיפק Google Translate, והמודל נדרש לשפרו לקסיקלית ותחבירית.
- ב. תרגום מלא מאנגלית לעברית - הפרומפט הכיל את משפט המקור באנגלית והמודל נדרש לתרגמו לעברית.

שיפור תרגום - שני המודלים הראו ביצועים לא מספיקים עבור מטלה (א), כאשר הבעיה העיקרית שעלתה מן הבדיקה הייתה הזיות (hallucinations) והוספת מידע שלא היה קיים במשפט המקור שניתן להם. לאחר הביצועים בבדיקה (א), הוחלט כי לא ניתן לסמוך על שיטה זו לתרגום בקנה מידה גדול הכולל מאות אלפי משפטים, וכי עבור הפרויקט המודלים יבצעו מטלת תרגום מלא מאנגלית לעברית (ב).

תרגום מלא באמצעות LLM - בוצעה השוואה מדגמית נוספת בת 100 משפטים בשפה האנגלית שתורגמו במלואם לעברית בעזרת Claude ו-Bard. איכות התרגום נבחנה דרך השוואת בלשנים (מעריכים אנושים).

שני הכלים הציגו ביצועים טובים יותר מ-Google Translate במטלת התרגום המלא, ועברו את הסף.

Claude - תרגום היטב 87 משפטים מתוך 100.

Bard - תרגום היטב 82 משפטים מתוך 100.

מן הבדיקה עלה כי Claude הציג יתרון קטן אל מול Bard (ב-5%).

בפברואר 2024, Google החליפה את מודל ה-AI שלה Bard במודל חדש ומשופר - Gemini. בדיקת תרגום מלא מאנגלית לעברית בוצעה עליו על מנת להיווכח אם הוא מספק תרגומים שווים בטיבם, או טובים יותר, מקודמו Bard. בבדיקה נמצא כי הוא זהה או טוב בביצועיו ל-Claude, ולעתים אף היה טוב ממנו מבחינת בחירה לקסיקלית (נוקשה פחות, ומדויקת יותר). לדוגמא:

משפט מקור (אנגלית)	תרגום Claude	תרגום Gemini
I never like to leave Canada, because I'm <b>disappointed</b> every time.	אני אף פעם לא אוהב לעזוב את קנדה, כי אני <b>מאכזב</b> כל פעם.	אני אף פעם לא אוהב לצאת מקנדה, כי כל פעם אני <b>מתאכזב</b> .
I beg your pardon. A man's voice beside her made her start and turn.	<b>אני מתנצל</b> . קול גברי לצדה גרם לה להיבהל ולהסתובב.	<b>סליחה</b> . קול גבר לידה גרם לה להיבהל ולהסתובב.

לאור תוצאות הבדיקות הוחלט לתרגם את כלל ה dataset בעזרת המודל של גוגל - Gemini. שני נושאים משמעותיים נוספים שעלו במהלך העבודה:

- א. ג'מיני לעתים נהג לחתוך משפטים באופן רנדומלי ולא להשלים את תרגומם - קושי זה נפתר ע"י הכנסת תו (@) בסוף כל משפט, והכנסת הוראה מפורשת בפרומפט לתרגם את המשפט עד תו זה.
- ב. ג'מיני סירב לתרגם משפטים המכילים תוכן מיני, תוכן אלים, ועלבונות וקללות - ומשפטים בעלי תכנים אלו לא נכנסו לדאטה-סט העברי.

שלב שלישי - תרגום ובדיקת כלל ה-dataset:

הדאטה בפרויקט ה-MNLI באנגלית הגיע מעשרה ז'אנרים (מקורות) שונים:

1. תמלול שיחות פנים אל פנים - *face-to-face*
2. דו"חות, נאומים והודעות לעיתונות מאתרי ממשלה - *Government*
3. מאגר מכתבים שנכתבו בשלהי שנות ה-90 ותחילת שנות ה-2000 - *Letters*
4. פרוטוקולים מתוך ועדה לחקירת מאורעות ה-11 בספטמבר - *Nine eleven*
5. פרסומים בנושא טקסטיל והתפתחות הילד בהוצאת אוני' אוקספורד - *OUP*
6. מאמרי תרבות פופולריים מן המגזין סלייט - *Slate*
7. תמלולי שיחות טלפון ממאגר של אוני' פנסילבניה - *Telephone*
8. מדריכי טיולים ותיירות בהוצאת ברליץ' - *Travel*
9. פוסטים קצרים בנושא בלשנות ושפות ממגזין וורבטיים - *Verbatim*
10. משפטים מיצירות פרוזה מודרניות - *Fiction*

לכל ז'אנר בוצעה בדיקה שכללה תרגום של משפטים אקראיים ובחינת תוצאות התרגום ע"י בלשנים. לאור תוצאות הבדיקה הוחלט שלא להשתמש בדאטה האנגלי המגיע מן המקורות המתומללים - *face-to-face*, ו-*telephone* לאור ההבנה כי כלים אוטומטיים אינם טובים בתרגום. המשפטים שנכללו במקורות הללו הכילו *filler words* למכביר כגון: *ummm, uh, like, you know, you see* וכו', משפטים מרובי משמעויות ללא הקשר (למשל משפט משיחה אנושית שתומלל בדאטה האנגלי כ-"*you know kids*" מתורגם ע"י כלים אוטומטיים לעברית רק כ-"אתה מכיר ילדים", בעוד שכוונת הדובר היא למשמעות "אתה יודע איך זה ילדים"), וכן משפטים שבורים - "*uh, yeah, maybe*" שהם חסרי ערך לפרויקט היסק לוגי בין משפטים.

הדאטה-סט של MNLI תורגם ע"פ ז'אנר (השמונה שנותרו מתוך העשרה לאחר ניפוי שני הז'אנרים המתומללים). לכל אחד מהז'אנרים נכתב פרומפט הכולל 5 דוגמאות לתרגום טוב מאנגלית לעברית, 2 ראשונות כלליות שחזרו בכל פרומפט, ו-3 דוגמאות לתרגום טוב למשפטים מתוך הז'אנר הספציפי.

להלן דוגמה לפרומפט לתרגום משפטים מתוך הוצאת אוניברסיטת אוקספורד Oup:

*English Origin: "One of our number will carry out your instructions minutely."*

*Hebrew Translation: "אחד מאיתנו יבצע את ההוראות שלך בדקדקנות."*

*English Origin: "Vrenna and I both fought him and he nearly took us."*

*Hebrew Translation: "ורנה ואני נלחמנו בו, והוא כמעט ניצח אותנו."*

*English Origin: "It has long been known that inventories at the macroeconomic level affect the depth and length of business cycles."*

*Hebrew Translation: "ידוע כבר זמן רב שמלאים ברמה המקרו-כלכלית משפיעים על עומק ואורך של מחזורי עסקים."*

*English Origin: "Only after a common convention for bar codes had been established and in use for several years did retailers turn to such systems to transmit and receive data from suppliers."*  
*Hebrew Translation: "רק לאחר שהוקמה קונבנציה משותפת לקודי ברקוד והייתה בשימוש במשך מספר שנים, קמעונאים פנו למערכות כאלה כדי להעביר ולקבל נתונים מספקים."*

*English Origin: "Children of highly involved fathers score better on measures of intelligence, school achievement, mature social behavior, and flexible beliefs about gender roles'in short, on all the positive outcomes associated with maternal employment."*  
*Hebrew Translation: "ילדים של אבות מעורבים במידה רבה, מציגים ציונים גבוהים יותר במבחני אינטליגנציה, הישגים לימודיים, התנהגות חברתית בשלה, ותפיסות גמישות יותר לגבי תפקידי מגדר בקיצור, בכל התוצאות החיוביות הקשורות לתעסוקת אמהות."*

*Translate these sentences from English to Hebrew up to the @. If there are human entities in the original sentence which are not explicitly female then assume they are male:*

היות ונתקלנו במקרים של כשלונות אקראיים לתרגם, הפרומפט הורץ בלולאה לתרגום 5 משפטים בכל פעם - כל קריאה לגימיני הכילה 5 משפטים לתרגום. בסוף תרגום כל ז'אנר נעשה נסיון נוסף לתרגם את כלל המשפטים שלא תורגמו כדי להבטיח שכל המשפטים יקבלו תרגום, בפעימות של 4 משפטים. אלו שעדיין לא תורגמו, נשלחו שוב בפעימות של 3 משפטים, וחוזר חלילה עד שכל המשפטים תורגמו. אם משפט חזר לא מתורגם גם לאחר שליחת קריאה איתו בלבד - הוא נותר חסר תרגום.

תרגום המשפטים בגימיני לכל ז'אנר חולק לשני שלבים - תרגום 1000 משפטים ראשונים לבדיקת איכות ראשונית, ותרגום כל הז'אנר:



לאחר הרצת אלף משפטים בנפרד לכל ז'אנר התקבלו התוצאות הבאות:

Genre	Size in Corpus	# of failed translation	% of Passable Translations
nine eleven	small	3.30%	96.67%
letters	small	0.000%	93.33%
fiction	large	2.803%	93.33%
government	large	0.000%	100.00%
slate	large	4.90%	83.33%
travel	large	0.200%	96.67%
verbatim	small	1.80%	93.33%
oup	small	0.000%	96.67%

בשלב זה הוחלט לבצע תרגום של כלל המשפטים מכל ז'אנרים.  
לאחר הרצת כל המשפטים בכל שמונת הז'אנרים (שערכה מספר ימים) התקבלו התוצאות הבאות:

Genre	Size in Corpus	# of Blanks	% of Passable Translations
nine eleven	small	3.852%	93.333%
letters	small	1.400%	93.333%

fiction	large	2.536%	90.323%
government	large	0.058%	96.667%
slate	large	4.114%	83.333%
travel	large	0.004%	90.000%
verbatim	small	5.853%	80.000%
oup	small	0.700%	86.667%

בשלב זה הוחלט לנפות את הז'אנר verbatim משתי סיבות - אחוז כשלונות תרגום גבוה, ותרגום לעברית ברמה לא טובה. ז'אנר זה נלקח ממגזין הדן בנושאי בלשנות ושפות בשפה האנגלית והוא דן בדקדוק ובנושאים נוספים ספציפיים לשפה האנגלית באופן שלעיתים לא עובר בתרגום לעברית. לדוגמא:

משפט מקור (אנגלית)	תרגום Gemini
Middle English <b>at + on</b> is the origin of Atone.	אנגלית תיכונה <b>ב- + ב-</b> היא המקור של Atone.
Therefore, statements referring <b>to giving me a shot</b> were comprehended.	לכן, הצהרות המתייחסות <b>למתן לי הזדמנות</b> הובנו.

היות ומדובר בתרגום אוטומטי, ולמרות שבוצעו אלפי בדיקות איכות, אנו מניחים שקיימים גם מקרים בהם התרגום אינו מיטבי.

## סיכום:

בעבודה זו מונגש לציבור דאטה-סט ראשון מסוג בעברית עבור מטלת NLI שהיא משימה חשובה בעיבוד שפה טבעית - להבנת קשרים ויחסים טקסטואלים, לבחינת ביצועי מודלי שפה גדולים ולצרכים נוספים. העבודה בוצעה דרך תרגום של dataset גדול בשפה האנגלית - multinli, מה שאפשר הנגשת כמות משפטים גדולה מאד. במסגרת העבודה נבחנו מס' פתרונות תרגום ע"י LLM, נמצאה ונבחנה שיטה מיטבית לתרגום משפטים ממספר זא'נרים שונים, תוך הקפדה על בדיקות איכות התרגום לאורך כל העבודה.

הדאטה-סט העברי החדש מכיל 303,383 משפטים בעברית, ממגוון מקורות, נושאים וסגנונות כתיבה על פי הפירוט בטבלה, ונמצא ב-Git שלנו (<https://github.com/NNLP-IL/HebNLI>) וב-[hugging face](https://huggingface.com)

Genre	Total # of sentences
nine eleven	1878



letters	1974
fiction	73734
government	76953
slate	71082
travel	75776
oup	1986

## נספחים:

:Google Gemini prompts המלאים עבור כל אחד משמונת הז'אנרים שהורצו ב-Google Gemini:

### **Fiction**

English Origin: "One of our number will carry out your instructions minutely."

Hebrew Translation: "אחד מאיתנו יבצע את ההוראות שלך בדקדקנות."

English Origin: "Vrenna and I both fought him and he nearly took us."

Hebrew Translation: "ורנה ואני נלחמנו בו, והוא כמעט ניצח אותנו."

English Origin: "Everybody was impressed, and one man's eyelid even ticked a little, because he didn't realize that the speaker of this witty comments wasn't Czarek, but one of the Fodder Brothers (a term coined later by triumphant Pytlasinski)."

Hebrew Translation: "כולם התרשמו, ועפעף של איש אחד אפילו ריצד קצת, כי הוא לא הבין שדובר ההערות (השנונות האלה לא היה צ'רק, אלא אחד מהאחים פודר (מונח שנטבע מאוחר יותר על ידי פיטלסינסקי המנצח).

English Origin: "I am to take it then that you decline to say where you were at the time that Mr. Mace positively recognized you as entering the shop to purchase strychnine?"

Hebrew Translation: "אני אמור להבין אם כך שאתה מסרב לומר היכן היית באותו הזמן שמר מייס זיהה אותך?"

English Origin: "After a name change to 'Puddle Skin Care' and a contract with a chain of make-up shops Zedhwora, the brand reached an exclusive status and overtook other skin care innovations - extract of quails' tonsils, essence of rutabaga and double C vitamin."

Hebrew Translation: "Zedhwora וחוזה עם רשת חנויות איפור 'Puddle Skin Care' -לאחר שינוי השם ל-"כפול C המותג הגיע למעמד אקסקלוסיבי ועקף חידושי טיפוח נוספים - תמצית שקדי שליו, תמצית רוטבגה וויטמין

Translate these sentences from English to Hebrew up to the @. If there are human entities in the original sentence which are not explicitly female then assume they are male:

### **Government**

English Origin: "One of our number will carry out your instructions minutely."

Hebrew Translation: "אחד מאיתנו יבצע את ההוראות שלך בדקדקנות."

English Origin: "Vrenna and I both fought him and he nearly took us."

Hebrew Translation: "ורנה ואני נלחמנו בו, והוא כמעט ניצח אותנו."

English Origin: "The program manager for the National Organic Program said that AMS knew that the rule would be controversial, so AMS decided to take advantage of IT's potential to facilitate the comment process, allowing comments to be provided via mail, fax, and email."

Hebrew Translation: "ידעה שהכלל יהיה שנוי AMS-מנהל התוכנית של התוכנית הלאומית אורגנית אמר ש" כדי להקל על תהליך ההערות, ולאפשר מתן הערות IT-החליטה לנצל את הפוטנציאל של ה AMS במחלוקת, ולכן בדואר, בפקס ובמייל."

English Origin: "During much of the 1990's the number of active boilermakers had been declining due to very low employment levels resulting from very low activity in the utility power plant construction business."

Hebrew Translation: "במהלך רוב שנות ה-90 מספר יצרני הדודים הפעילים ירד עקב רמות תעסוקה נמוכות" מאוד כתוצאה מפעילות נמוכה מאוד בעסקי בניית תחנות כוח."

English Origin: "Also, much of the information maintained by federal agencies, although unclassified, is extremely sensitive, and many automated operations are attractive targets for individuals or organizations with malicious intentions, such as committing fraud for personal gain or sabotaging federal operations."

Hebrew Translation: "כמו כן, חלק ניכר מהמידע שנשמר על ידי סוכנויות פדרליות, למרות שאינו מסווג, הוא" רגיש ביותר, ופעולות אוטומטיות רבות הן מטרות אטרקטיביות עבור אנשים או ארגונים עם כוונות זדוניות, כגון ביצוע הונאה למטרות רווח אישי או חבלה בפעולות פדרליות."

Translate these sentences from English to Hebrew up to the @. If there are human entities in the original sentence which are not explicitly female then assume they are male:

### **Slate**

English Origin: "One of our number will carry out your instructions minutely."

Hebrew Translation: "אחד מאיתנו יבצע את ההוראות שלך בדקדקנות."

English Origin: "Vrenna and I both fought him and he nearly took us."

Hebrew Translation: "ורנה ואני נלחמנו בו, והוא כמעט ניצח אותנו."

English Origin: "For Klein, control is always bad--whether it's the evil health-and-fashion industry dictating our ideas about beauty, or the beleaguered dieter struggling to curb his cravings in a spirit of misguided self-hatred."

Hebrew Translation: "עבור קליין, שליטה היא תמיד גרועה - בין אם זו תעשיות הבריאות והאופנה המרושעות" שמכתיבות את הרעיונות שלנו לגבי יופי, או הדיאטן הנצור שנאבק לרסן את תשוקותיו ברוח של שגאה עצמית מוטעית."

English Origin: "All the avant-garde artists included are actually academic, both in the sense of deriving their ideas about cultural difference from French literary critics and in the sense that they follow the dictates of others about what art should be."

Hebrew Translation: "כל האמנים האוונגרדיים שנכללו הם למעשה אקדמיים, הן במובן של גזירת רעיונותיהם על" שנוי תרבותי ממבקרי ספרות צרפתים והן במובן שהם פועלים לפי תכתיבים של אחרים לגבי מה אמנות צריכה להיות."

English Origin: "For example, while we have turned our bathrooms into palaces of comfort, lots of the world's people still squat over holes, which makes it difficult to finish reading the business section, but is a real bone builder."

Hebrew Translation: "לדוגמה, בעוד שהפכנו את חדרי האמבטיה שלנו לארמונות של נוחות, הרבה מאנשי" העולם עדיין כורעים מעל חורים, מה שמקשה על סיום קריאת המדור העסקי, אבל בונה עצמות משמעותי."

Translate these sentences from English to Hebrew up to the @. If there are human entities in the original sentence which are not explicitly female then assume they are male:

## **Travel**

English Origin: "One of our number will carry out your instructions minutely."

Hebrew Translation: "אחד מאיתנו יבצע את ההוראות שלך בדקדקנות."

English Origin: "Vrenna and I both fought him and he nearly took us."

Hebrew Translation: "ורנה ואני נלחמנו בו, והוא כמעט ניצח אותנו."

English Origin: "This imposing structure, built by El-Jezzar, nowadays holds the Museum of the Underground Prisoners, which documents the Jewish resistance movement that fought against the British during the Mandate period (part of the film Exodus was shot here)."

Hebrew Translation: "המבנה המרשים הזה, שנבנה על ידי אל-ג'זאר, מכיל כיום את מוזיאון אסירי המחתרות, (המתעד את תנועת ההתנגדות היהודית שלחמה נגד הבריטים בתקופת המנדט (חלק מהסרט אקסודוס צולם כאן).

English Origin: "It has beautiful natural attractions nearby, and it makes a good base for other, longer journeys around Jamaica, being within relatively easy reach of Kingston, the Blue Mountains, and the coast road that leads to both Montego Bay and Port Antonio."

Hebrew Translation: "יש לו אטרקציות טבעיות יפות בקרבת מקום, והוא מהווה בסיס טוב לנסיעות אחרות, ארוכות יותר ברחבי ג'מייקה, בהיותו בקרבה יחסית לקינגסטון, להרים הכחולים ולכביש החוף שמוביל גם למונטגו וביי וגם לפורט אנטוניו."

English Origin: "The British Residency today preserved as a monument initially commemorated British resistance, but since Independence it has been visited by Indians interested in this relic of their own struggle for self-assertion, or interested in a family picnic on the lawns."

Hebrew Translation: "בית המגורים הבריטי השתמר היום כאנדרטה לזכר ההתנגדות הבריטית בתחילה, אך מאז עצמאותה ביקרו בו הודים המעוניינים בשריד זה של מאבקם העצמי, או מתעניינים בפיקניק משפחתי על המדשאות."

Translate these sentences from English to Hebrew up to the @. If there are human entities in the original sentence which are not explicitly female then assume they are male:

## **Nine Eleven**

English Origin: "One of our number will carry out your instructions minutely."

Hebrew Translation: "אחד מאיתנו יבצע את ההוראות שלך בדקדקנות."

English Origin: "Vrenna and I both fought him and he nearly took us."

Hebrew Translation: "ורנה ואני נלחמנו בו, והוא כמעט ניצח אותנו."

English Origin: "It sounds like, we're talking to New York, that there's another one aimed at the World Trade Center."

Hebrew Translation: "נשמע כאילו, אנחנו מדברים עם ניו יורק, שיש אחד נוסף המכוון אל מרכז הסחר העולמי."

English Origin: "At 7:15, a pair of them, Khalid al Mihdhar and Majed Moqed, checked in at the American Airlines ticket counter for Flight 77, bound for Los Angeles."

Hebrew Translation: "בשעה 7:15, זוג מהם, חאלד אל-מחדהאר ומאג'ד מוקד, הצטרפו בדלפק כרטיסי הטיסה של חברת התעופה אמריקן איירליינס לטיסה 77 לכיוון לוס אנג'לס."

English Origin: "The long, disjointed document condemned the Saudi monarchy for allowing the presence of an army of infidels in a land with the sites most sacred to Islam, and celebrated recent suicide bombings of American military facilities in the Kingdom."

Hebrew Translation: "המסמך הארוך והלא רציף גינה את המלוכה הסעודית על שהתירה נוכחות של צבא של כופרים בארץ עם האתרים הקדושים ביותר לאסלאם, וחגג את פיגועי ההתאבדות האחרונים במתקנים צבאיים באמריקה בממלכה."

Translate these sentences from English to Hebrew up to the @. If there are human entities in the original sentence which are not explicitly female then assume they are male:

### **Letters**

English Origin: "One of our number will carry out your instructions minutely."

Hebrew Translation: "אחד מאיתנו יבצע את ההוראות שלך בדקדקנות."

English Origin: "Vrenna and I both fought him and he nearly took us."

Hebrew Translation: "ורנה ואני נלחמנו בו, והוא כמעט ניצח אותנו."

English Origin: "In practice, however, retail inventory management is fraught with challenges, such as long and uncertain order-fulfillment lead times, and errors in product identification and record keeping."

Hebrew Translation: "בפועל, אולם, ניהול המלאי בקמעונאות רצוף אתגרים, כמו זמני המתנה ארוכים וחוסר וודאות למילוי הזמנות, ושגיאות בדיהוי מוצרים ובשמירת רשומות."

English Origin: "It is a pivotal time in the life of a young girl-when small changes in her behavior, or the choices she makes, can affect her future."

Hebrew Translation: "זהו זמן משמעותי בחייה של ילדה צעירה - כאשר שינויים קטנים בהתנהגות שלה, או הבחירות שהיא עושה, יכולים להשפיע על עתידה."

English Origin: "In addition, gifts provide essential support in a number of other areas, including recruitment of new faculty members and students, maintaining our Career Services office, securing guest speakers, supporting the school's commencement and publications"

Hebrew Translation: "בנוסף, מתנות מספקות תמיכה חיונית בכמה תחומים אחרים, כולל גיוס חברי סגל ותלמידים חדשים, תחזוקת משרד שירותי הקריירה שלנו, השגת דוברים אורחים, תמיכה באירועי הסיום והפרסומים של בית הספר."

Translate these sentences from English to Hebrew up to the @. If there are human entities in the original sentence which are not explicitly female then assume they are male:

### **Oup (Oxford University Press)**

English Origin: "One of our number will carry out your instructions minutely."

Hebrew Translation: "אחד מאיתנו יבצע את ההוראות שלך בדקדקנות."

English Origin: "Vrenna and I both fought him and he nearly took us."

Hebrew Translation: "ורנה ואני נלחמנו בו, והוא כמעט ניצח אותנו."

English Origin: "It has long been known that inventories at the macroeconomic level affect the depth and length of business cycles."

Hebrew Translation: "ידוע כבר זמן רב שמלאים ברמה המקרו-כלכלית משפיעים על עומק ואורך של מחזורי עסקים."

English Origin: "Only after a common convention for bar codes had been established and in use for several years did retailers turn to such systems to transmit and receive data from suppliers."

Hebrew Translation: "רק לאחר שהוקמה קונבנציה משותפת לקודי ברקוד והייתה בשימוש במשך מספר שנים, רק לאחר שהוקמה קונבנציה משותפת לקודי ברקוד והייתה בשימוש במשך מספר שנים, פנו למערכות כאלה כדי להעביר ולקבל נתונים מספקים."

English Origin: "Children of highly involved fathers score better on measures of intelligence, school achievement, mature social behavior, and flexible beliefs about gender roles'in short, on all the positive outcomes associated with maternal employment."

Hebrew Translation: "ילדים של אבות מעורבים במידה רבה, מציגים ציונים גבוהים יותר במבחני אינטליגנציה, הישגים לימודיים, התנהגות חברתית בשלה, ותפיסות גמישות יותר לגבי תפקידי מגדר בקיצור, בכל התוצאות החיוביות הקשורות לתעסוקת אמהות."

Translate these sentences from English to Hebrew up to the @. If there are human entities in the original sentence which are not explicitly female then assume they are male:

### **Verbatim**

English Origin: "One of our number will carry out your instructions minutely."

Hebrew Translation: "אחד מאיתנו יבצע את ההוראות שלך בדקדקנות."

English Origin: "Vrenna and I both fought him and he nearly took us."

Hebrew Translation: "ורנה ואני נלחמנו בו, והוא כמעט ניצח אותנו."

English Origin: "Thus the evidence at the RHD offices indicated the need for reinforcement of the notion of maleness associated with the word, not, as M and S suggest, a trend toward epicenism."

Hebrew Translation: "הצביעו על הצורך לחזק את התפיסה של זכריות RHD-לפיכך, הראיות במשרדי ה' הקשורה למילה, ולא, כפי שמ' וס' מציעים, מגמה לעבר שוויון מגדרי."

English Origin: "While writers whose business it is to be witty often fail to produce, grave authors occasionally are mirthful when laughter is farthest from their minds."

Hebrew Translation: "בעוד שסופרים שעסקם הוא להיות שנונים לעתים קרובות נכשלים בכך, מחברים רציניים לפעמים מצחיקים באופן בלתי צפוי כשצחוק הוא הדבר הרחוק ביותר ממחשבותיהם."

English Origin: "But the power of words is such that I contend their impact is totally lost if not treated somehow, there is a difference between Someone had written shit on the wall and Some asshole had written shit on the wall : the writer of the former has more credibility; the writer of the latter was, very likely, the very asshole who had written shit on the wall."

Hebrew Translation: "אבל כוחם של מילים הוא כזה שאני טוען שהשפעתם נעלמת לחלוטין אם הן לא מטופלות" איכשהו, יש הבדל בין "מישהו כתב חרא על הקיר" ל"איזה חור תחת כתב חרא על הקיר": הכותב של האפשרות הראשונה נהנה ממהימנות רבה יותר; הכותב של האפשרות השנייה, היה כנראה, אותו חור תחת שכתב חרא על הקיר."

Translate these sentences from English to Hebrew up to the @. If there are human entities in the original sentence which are not explicitly female then assume they are male: