

התרומה הדיפרנציאלית של מאפייני טקסט להערכה ממוחשבת של חיבורים מסוגים שונים

ענת בן-סימון

פרויקט השפה העברית (HLP), מאל"ו

הכנס השביעי של אפי, 2011, ירושלים

מטרת המחקר

◆ זיהוי מאפייני הטקסט המנבאים באופן הטוב ביותר את איכות ההבעה בכתב (ציוני חיבור) של מעריכים מומחים.

שיטת המחקר: מדגם

- ◆ שני מקבצים של חיבורים שכתבו תלמידי כיתה י"ב בשיעורי הבעה בכתב (מתוך עבודת הדוקטורט של עירית השכל-שחם, 2010)
- ◆ שני מקבצים של חיבורים שכתבו מועמדים למוסדות להשכלה גבוהה במסגרת מבחן יע"ל (מבחן ידע בעברית)

מבחן יע"ל דוברי עברית שפה שנייה		תלמידי י"ב דוברי עברית שפת אם		
2009	2001	בתר	קדם	
496	484	294	368	N
980		662		

שיטת המחקר: ניתוח הנתונים

◆ לכל חיבור בארבעת המדגמים הופקו 61 מאפיינים סטטיסטיים, מורפולוגיים, לקסיקליים ומורפו-תחביריים.

◆ בכל מדגם חושבו המאפיינים בשלוש רמות קיבוץ -

◆ 61 מאפיינים בדידים

◆ 38 מאפיינים מקובצים

◆ 15 גורמים

◆ לכל מדגם ורמת קיבוץ של מאפיינים –

◆ הופקו משוואות רגרסיה לניבוי הציונים

◆ נעשה תיקוף צולב לבדיקת יציבות תוקף הניבוי

◆ הקריטריון לניבוי - ממוצע הציונים בחיבור של שני מעריכים

תוצאות

מבחן יע"ל דוברי עברית שפה שנייה		תלמידי י"ב דוברי עברית שפת אם		
2009	2001	בתר	קדם	מתאם
.83	.83	.76	.68	מחשב / הערכות מומחים (משקלות אופטימליים)
.83		.75		
.80	.93	.81	.72	מהימנות בין שופטים
.88		.80		
496	484	294	368	N

תיקוף צולב לפי גודל מדגם אימון

חיבורי תלמידי י"ב (עברית שפת אם)

מתאם מרובה	80% מהמדגם	50% מהמדגם	20% מהמדגם	
משקלות אופטימליים	טווח	.73-.78	.73-.80	.73-.78
	ממוצע	.75	.74	.76
תיקוף צולב	טווח	.72-.78	.71-.76	.69-.73
	ממוצע	.76	.73	.71
מתאם בין שופטים		.80		

הניתוח מבוסס על 661 חיבורים ו- 38 מאפיינים

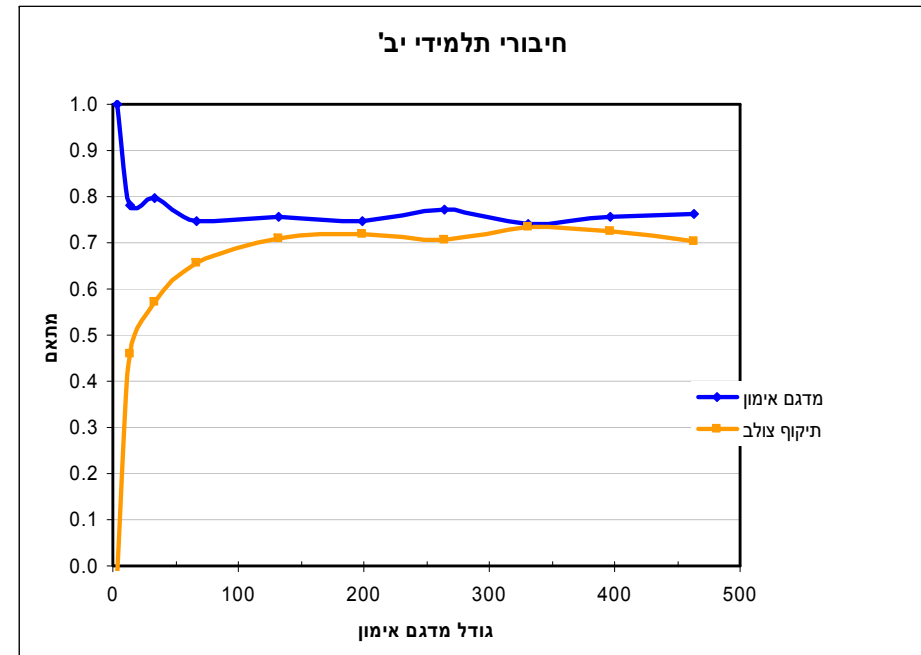
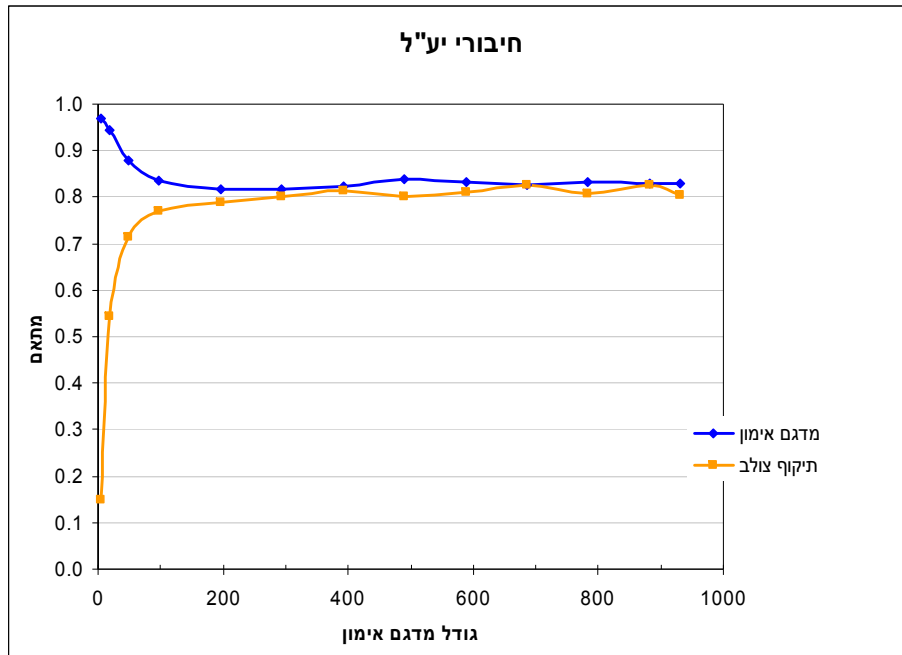
תיקוף צולב לפי גודל מדגם אימון

חיבורי יע"ל (עברית שפה זרה)

מתאם מרובה	80% מהמדגם	50% מהמדגם	20% מהמדגם	
משקלות אופטימליים	טווח	.82-.85	.81-.86	.79-.84
	ממוצע	.83	.84	.82
תיקוף צולב	טווח	.75-.83	.78-.83	.60-.80
	ממוצע	.81	.80	.79
מתאם בין שופטים		.88		

הניתוח מבוסס על 980 חיבורים ו- 38 מאפיינים

תיקוף צולב לפי גודל מדגם אימון



ניבוי ציון לפי רמות קיבוץ של מאפיינים

חיבורי תלמידי י"ב (עברית שפת אם)

15 גורמים	38 מאפיינים מקובצים	61 מאפיינים		
.74	.74	.77	מתאם ממוצע	משקלות אופטימליים
6	6	8	מס' מאפיינים	
.74	.73	.72	מתאם ממוצע	תיקוף צולב מדגם חצוי (50%)
.80			מתאם בין שופטים	

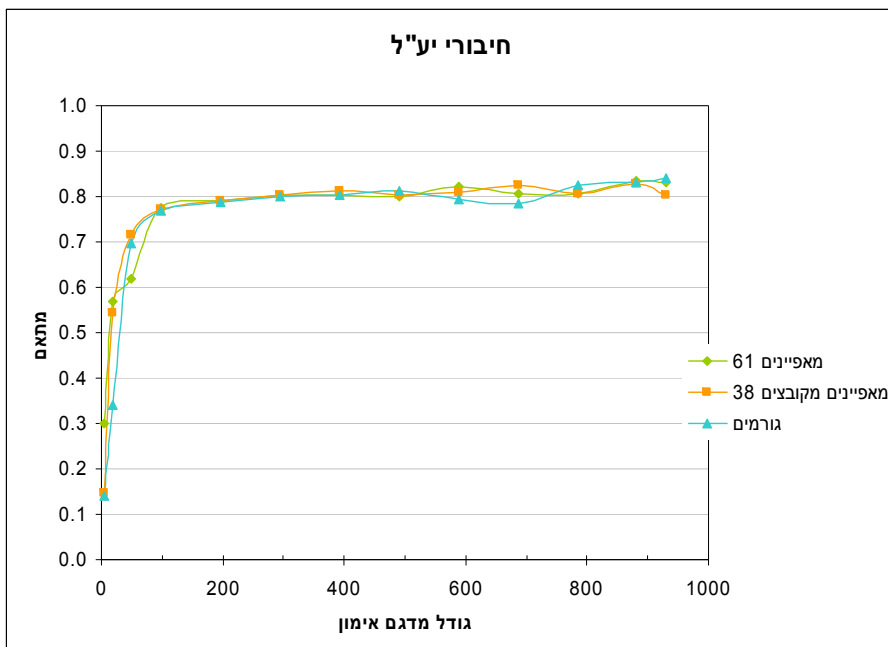
ניבוי ציון לפי רמות קיבוץ של מאפיינים

חיבורי יע"ל (עברית שפה זרה)

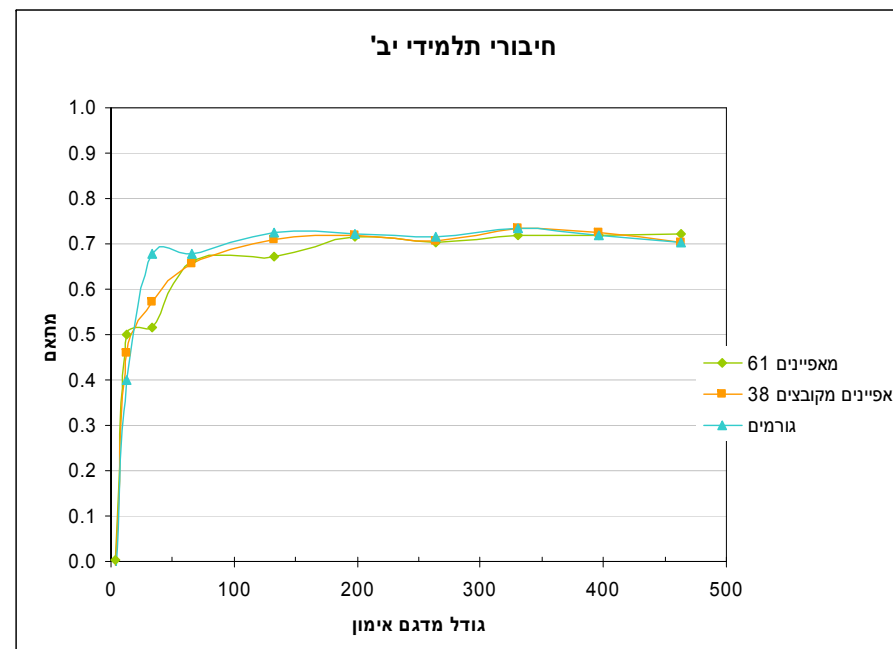
15 גורמים	38 מאפיינים מקובצים	61 מאפיינים		
.81	.84	.85	מתאם ממוצע	משקלות אופטימליים
13	12	17	מס' מאפיינים	
.81	.80	.80	מתאם ממוצע	תיקוף צולב (מדגם חצוי (50%))
.88			מתאם בין שופטים	

דיוק הניבוי (תיקוף צולב) לפי רמת הקיבוץ של המאפיינים וגודל מדגם האימון

חיבורי יע"ל



חיבורי תלמידי יב'



מאפייני טקסט 'מצטיינים'

מתאמים עם הערכות מומחים

טקסטים נרטיביים	חיבורי יע"ל		חיבורי י"ב		מאפיין	תת- גורם
	2009	2001	בתר	קדם		
.79	.67	.75	.72	.67	גיוון לקסיקלי	2.2
.78	.49	.70	.69	.64	מס' מחרוזת	2.1
.53	.38	.51	.47	.44	גיוון מילות יחס	2.4
-.79	-.47	-.33	-.59	-.44	שכיחות מילים בשפה	2.3
.70	.38	-.30	.39	.35	גיוון לשוני (TTR_D)	2.6
.30	.29	.36	.35	.29	גיוון סימני פיסוק	6.1
.29	.21	.40	.35	.24	גיוון מילות חיבור וקישור	7.1
.56	.50	.53	.39	.22	גיוון בניינים	15.1
-.22	.26	.36	.20	.19	גיוון סימני פיסוק	6.1

מאפייני טקסט "מצטיינים"

(עברית שפה זרה)

טקסטים נרטיביים	חיבורי יע"ל		חיבורי י"ב		מאפיין	תת- גורם
	2009	2001	בתר	קדם		
.22	.31	.33	.24	.13	% שמות תואר	1.3
.22	.25	.26	.24	.12	% מילות יחס	10.2
.03	.29	.35	.21	.07	דחיסות תוכנית: % מילות תוכן	1.2
-.06	-.37	-.39	-.26	-.10	דחיסות תוכנית: % המחרוזות השכיחות	1.2
.40	.43	.52	.09	.01	אורך ממוצע של מחרוזת	1.1
.37	.27	.26	.16	.01	% בניין 'התפעל'	13.1
.05	-.35	-.30	-.10	-.09	% מילים עם שגיאות כתיב	11.2

סיכום

◆ **בניתוח התוצאות נמצאו מאפיינים רבים העומדים במתאם גבוה עם הציונים של איכות הבעה בכתב. הבולטים בהם הם:**

◆ אורך הטקסט (.70-.49)

◆ גיוון לקסיקלי (.75-.67)

◆ שכיחות המילים בשפה (.59-.33)

◆ גיוון השימוש במילות יחס (.51-.38)

◆ גיוון השימוש בסימני פיסוק (.36-.29)

◆ **כמה מהמאפיינים נמצאו רלוונטיים להערכת הבעה בכתב בעברית כשפה זרה בלבד, למשל:**

◆ אורך ממוצע של מחרוזת (.53-.42)

◆ שכיחות יחסית של שגיאות איות (.35-.30)

סיכום (המשך)

◆ בהתייחס לניבוי הציונים באמצעות משוואת רגרסיה נמצא כי ניתן להסביר כ-

◆ 56% משונות הציונים בחיבורים של תלמידי כיתה י"ב ילידי הארץ
($R=0.75$).

◆ 69% משונות הציונים בחיבורים של נבחני יע"ל, דוברי עברית שפה זרה
($R=0.83$)

◆ בשני המדגמים נמצאו פערים קטנים (כ-0.04) בין המתאם המרובה שהתבסס על משקלות אופטימאליים לבין התיקוף הצולב

◆ לאופן קיבוץ המאפיינים הייתה השפעה מזערית על דיוק הניבוי והתיקוף הצולב

◆ התיקוף הצולב מושפע מגודל מדגם האימון, עם זאת דיוק הניבוי מתייצב במדגמים בני 150-200 חיבורים בלבד

ומה בהמשך ...

- ◆ תיקוף המאפיינים מול סולמות הערכה (למשל, תוכן מבנה לשון)
- ◆ פיתוח מאפיינים נוספים להערכת מורכבות תחבירית, תקינות דקדוקית, ועושר תכני ורעיוני

תודה לצוות הפרויקט

ותודה לקהל

