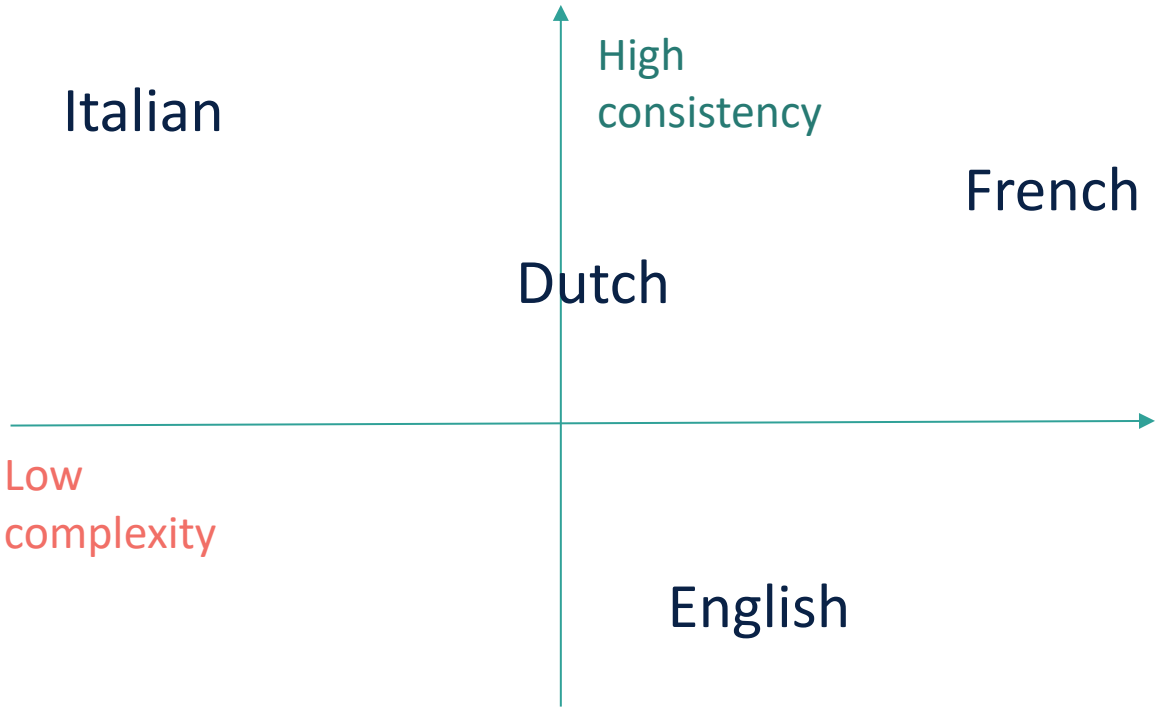
Rethinking morphology-based reading in Hebrew: new findings using a finger-tracking paradigm

Naama Evanhaim, Amalia Bar-On

Tel Aviv University
Israel

Writing systems in European orthographies differ from one another in their orthographic depth



Schmalz, Marinus, Coltheart & Castles, 2015

The Hebrew writing system comprises two versions

Transparent orthography = a pointed version (mostly serves novice readers)

Nearly complete phonological representation:

✓ Consonants

✓ Vowel letters

✓ Diacritical marks (*nikud*; Hebrew ‘points’)

מִקְלֶדֶת

miklédet

[keyboard]

Katz & Frost, 1992; Ravid, 2005; Shimron, 2003

The Hebrew writing system comprises two versions

Opaque orthography = an unpointed version (serves Hebrew readers from relatively early stages onward)

- ✓ Discards the diacritical marks
- ✓ Partial and ambiguous phonological representation
- ✓ Essentially less alphabetic and more abjad-like

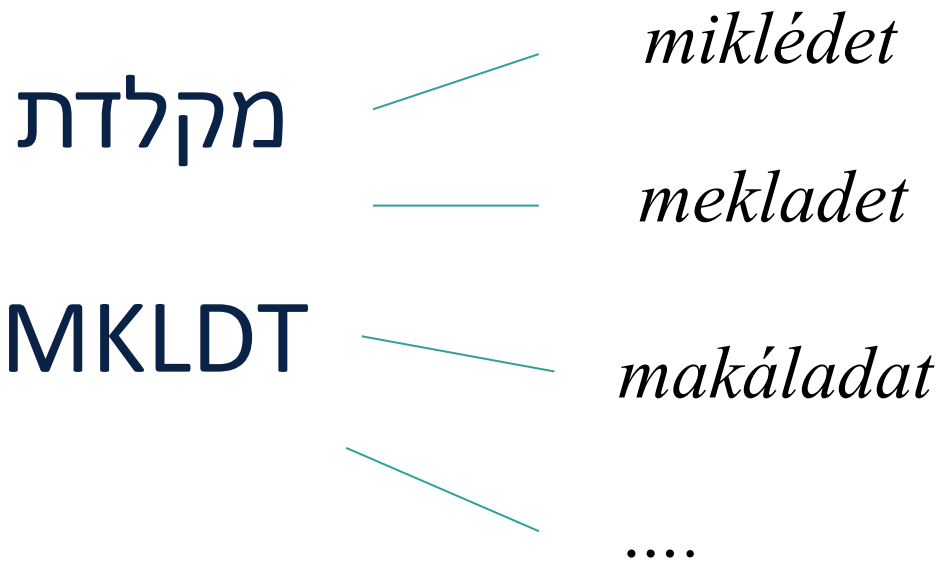
מקלדת

miklédet

[keyboard]

Katz & Frost, 1992; Ravid, 2005; Shimron, 2003

Abjadic orthographies, like Hebrew and Arabic, present a different aspect of orthographic depth termed incompleteness presenting a **qualitatively different** problem



Katz & Frost, 1992; Ravid, 2005; Shimron, 2003; Schmaltz et al., 2015

Hebrew's rich morphological structure compensates for the orthographic incompleteness through morphological transparency

Root-and-pattern morphological structure

miklédet ***[keyboard]*** ***מקלדת***

pattern

miCCéCet

Root

k-l-d ק-ל-ד

מקלדת = מ888ת

Morpho-orthographic structure

Bar-On & Ravid, 2011, Deutsch & Frost, 2003, Deutsch and Kuperman, 2019, Frost, Deutsch & Forester, 2000, Share & Bar-On, 2018, Velan & Michaly, 2016

miklédet [keyboard] מקלדת

mirpéset [balcony] מרפסת

miktéret [pipe] מקטרת

mišméret [shift] משמרת

מ888ת

No vowel representation

tikšóret [communication] תקשורת

tizmóret [orchestra] תזמורת

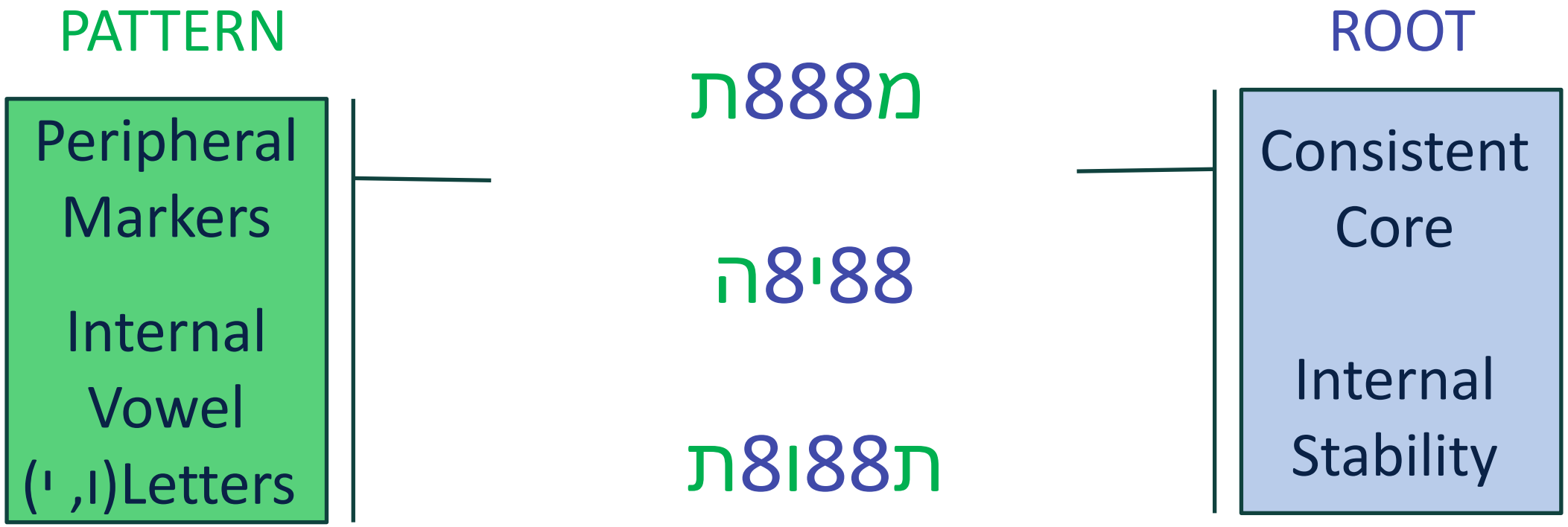
šmira [guarding] שמירה

bdika [testing] בדיקה

ת88ו8ת 88י8ה

Partial and opaque representation by the four vowel letters אהו"י AHWY

Morphological transparency



Structure highlights root consistency versus peripheral pattern indicators and internal vowel letters

Hebrew readers establish the ability to rely on **morpho-orthographic cues** for recovering the missing phonological information and reaching successful **morpho-orthographic identification**

מ888ת 88י8ה ת88ו8ת

Bar-On & Ravid 2011; Bar-On, 2015

Morpho-orthographic identification

More accurate reading as early as second grade

Faster reading in second, fourth, and sixth-grade students

morphological decomposition into its root and pattern morphemes

Bar-On, Shalhoub-Awwad & Tuma-Basila, 2018; Bitan, Weiss, Katzir, and Truzman, 2020 ; Deutsch & Frost, 2003; Deutsch, Velan & Michaly, 2016; Frost, Deutsch & Forester, 2000; Norman, Degani & Peleg, 2017; Shany, Bar-On, and Katzir, 2012

Morpho-orthographic identification

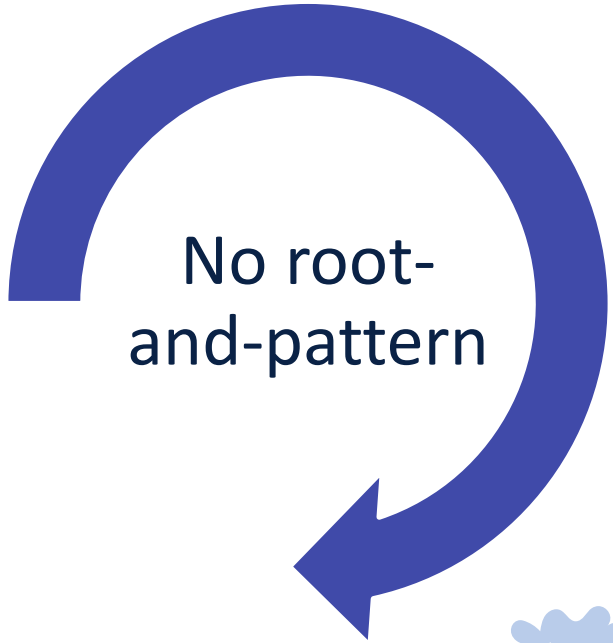
Root-and-pattern nature of Semitic content words

But…

The Morphological Gap

✓ Verbs: Fully root-and-pattern based

✗ Nouns & Adjectives: Many exceptions



- **Loanwords**:

feminizem [feminism] פמיניזם ; *universita* [university] אוניברסיטה

- **Non-morphological native Hebrew words:**

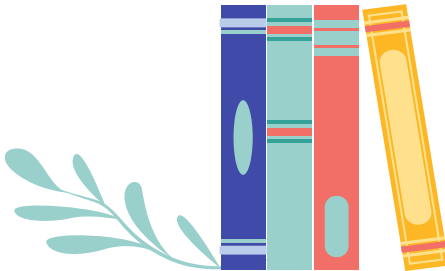
yad [hand] יד ; *agvanya* [tomato] עגבניה

Morphological Structure ≠ Orthographic Transparency

Root-and-pattern morphological structure

does not automatically guarantee efficient

morpho-orthographic identification



Morphological Structure ≠ Orthographic Transparency

A continuum of pattern cue availability

tiCCóCet *miCCéCet* *CCóCet* *miCCaC* *CiCaC* *CéCeC*

תסוסת מססת סוסת מסס סיס סס

Two pattern letters → Pattern letter+ה → One pattern letter → Only final ה → No pattern letters

מססה מסיה ססה סיה

Morphological Structure ≠ Orthographic Transparency

A continuum of pattern cue availability

	meCaCéCet		*maCCeC*		*CaCeC*
					CaCaC
tiCCóCet	*miCCéCet*	*CCóCet*	*miCCaC*	*CiCaC*	*CéCeC*
ת88ו8ת	מ888ת	88ו8ת	מ888	8י88	888

Homography:

Same structure → Multiple patterns

A Widespread Phenomenon

Morphological Structure ≠ Orthographic Transparency

Allomorphic variants:

Root-pattern boundaries become opaque

due to defective roots

Challenging root and pattern distinction

mapal [waterfall] מפל

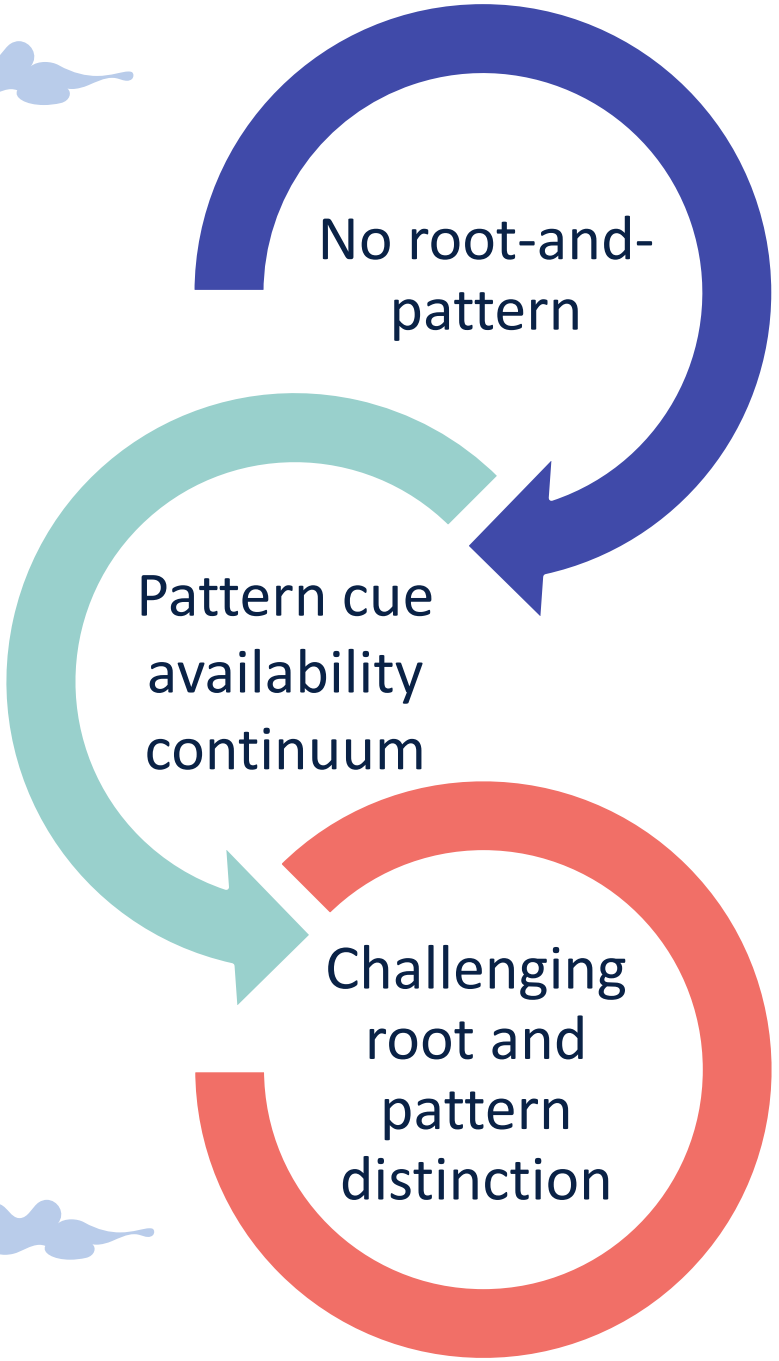
[Original pattern: *miCCaC*]

[Morpho-orthographic structure: 88מ]

ganan [gardener] גנן

[Original pattern: *CaCaC*]

[Morpho-orthographic structure: 888]



✓ Structure provides valuable cues

BUT:

✗ Does not guarantee efficient identification

✗ Root-pattern distinction can be challenging

✗ Non-negligible group of non-morphological words

Corpus Study

Evanhaim & Bar-On, in preparation

Despite the wealth of information on Hebrew morphological structures

There is still limited information on the extent to which these characteristics distribute in texts addressed to readers

Corpus Study

Evanhaim & Bar-On, in preparation

- ✓ Nouns from early reading texts
  (1st – 4th grade; 239 texts, 15,108 words, 5,824 nouns)

  Largest lexical category ;

  Shares patterns with adjectives ;

  Includes non-root-and-pattern forms

Corpus Study

Evanhaim & Bar-On, in preparation

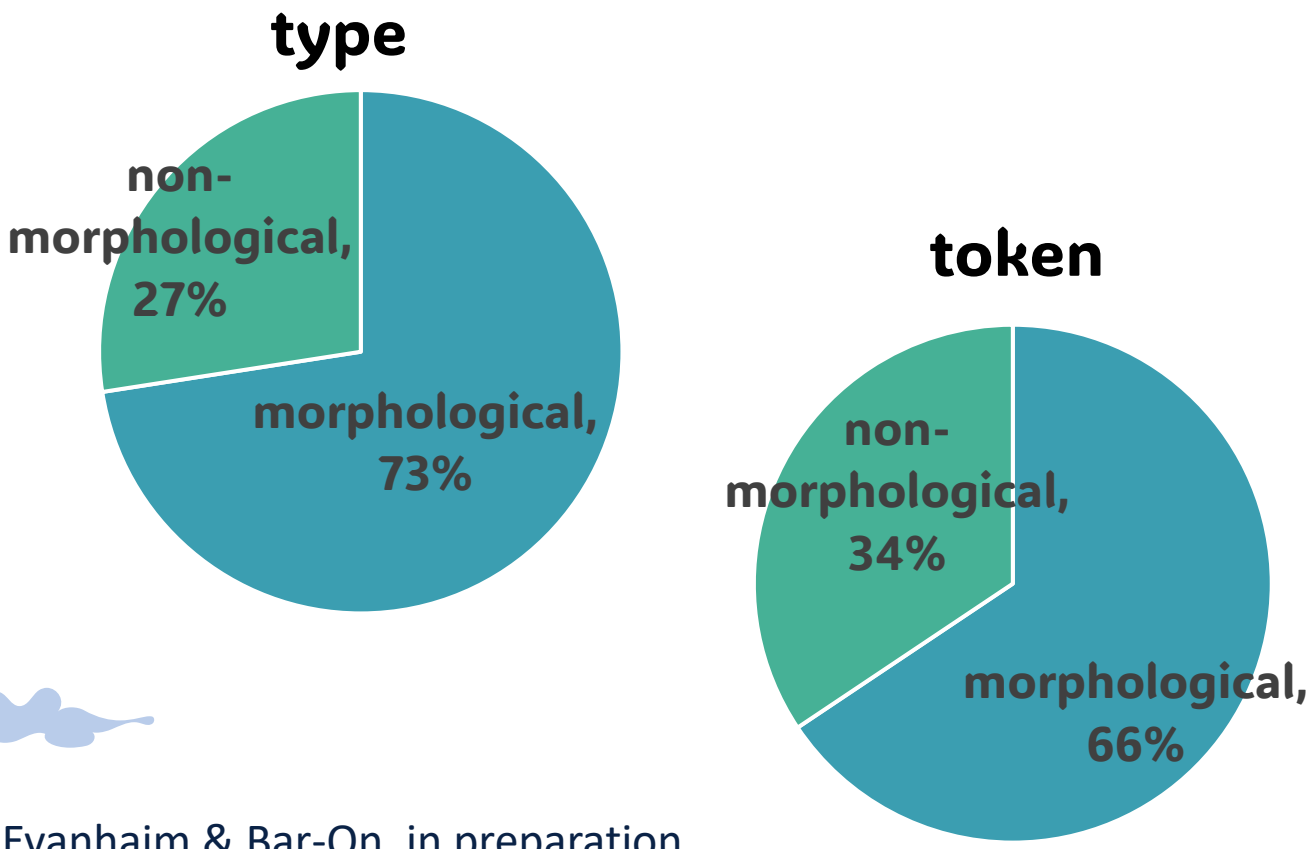
- ✓ Nouns from early reading texts
- ✓ Morphological and morpho-orthographic structures of noun lemmas
- ✓ Noun structure mapping - Phonological-orthographic coding and analysis

Infers on nouns' morphological patterns instead of looking for them a priori

Interested in the structure the reader encounters
(did not mark nouns origin pattern)

Corpus Study | Findings

Morphological vs. non-morphological-based structures



Orthographic Density

1344 noun types

253 vowel structures

201 orthographic and morpho-orthographic structures

Evanhaim & Bar-On, in preparation

Corpus Study | Findings

Non-morphological-based nouns [27%]

→ no morpho-orthographic Support yet..

Idiosyncratic nouns [10%]

Nouns with repetitive vowel structure [17%]

dinosaur
דינוזאור ‘dinosaur’

afarsek
אפרסק ‘peach’

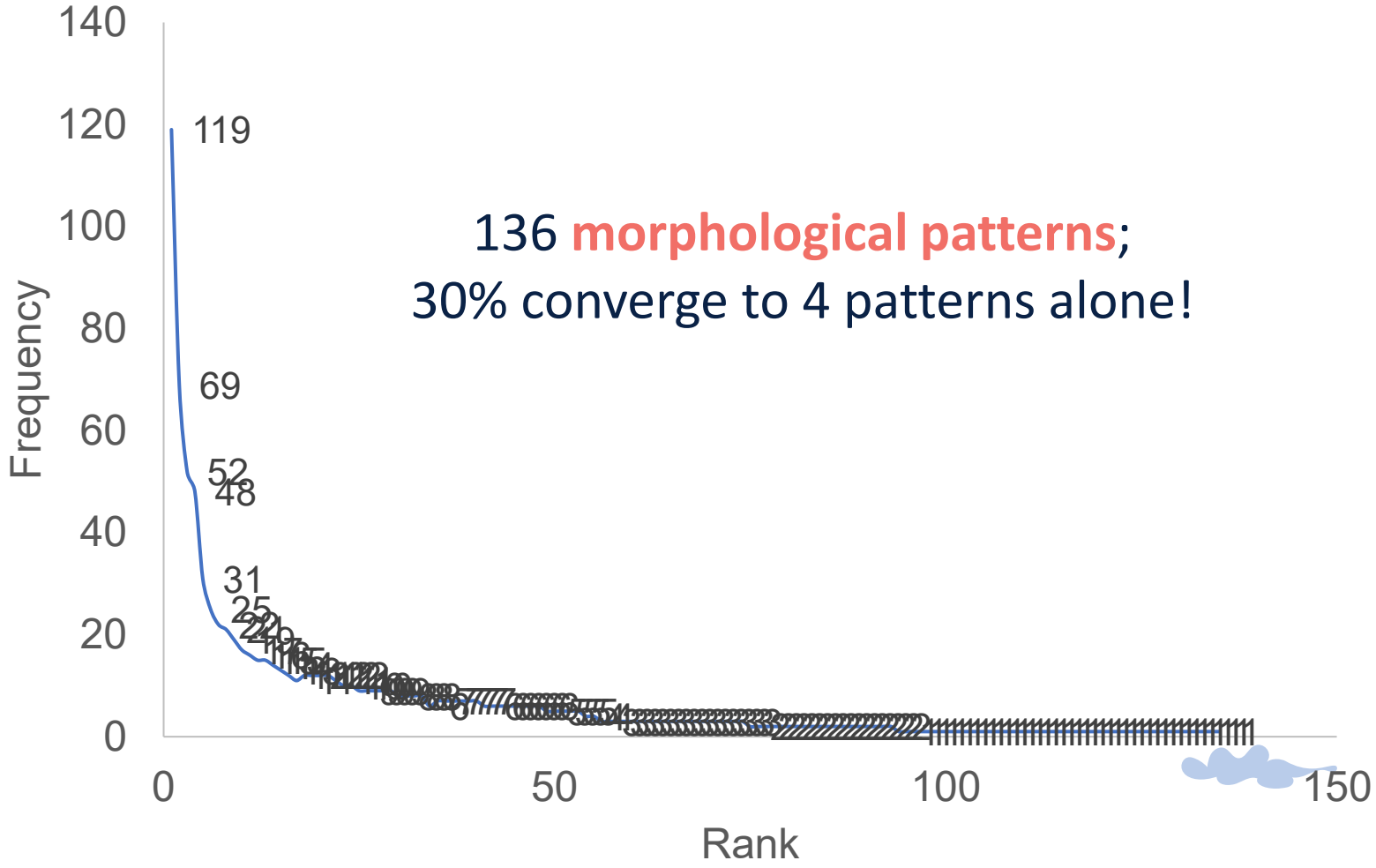
- ✓ High level of repetitiveness - 228 noun types converged into 27 phonological structures alone
- ✓ Mostly Simple vowel structure (60% monosyllabic)
- ✓ Most nouns are concrete (basic semantic-lexical level)

Evanhaim & Bar-On, in preparation

Morphological Density

944 noun types

136 morphological patterns



136 **morphological patterns**;
30% converge to 4 patterns alone!

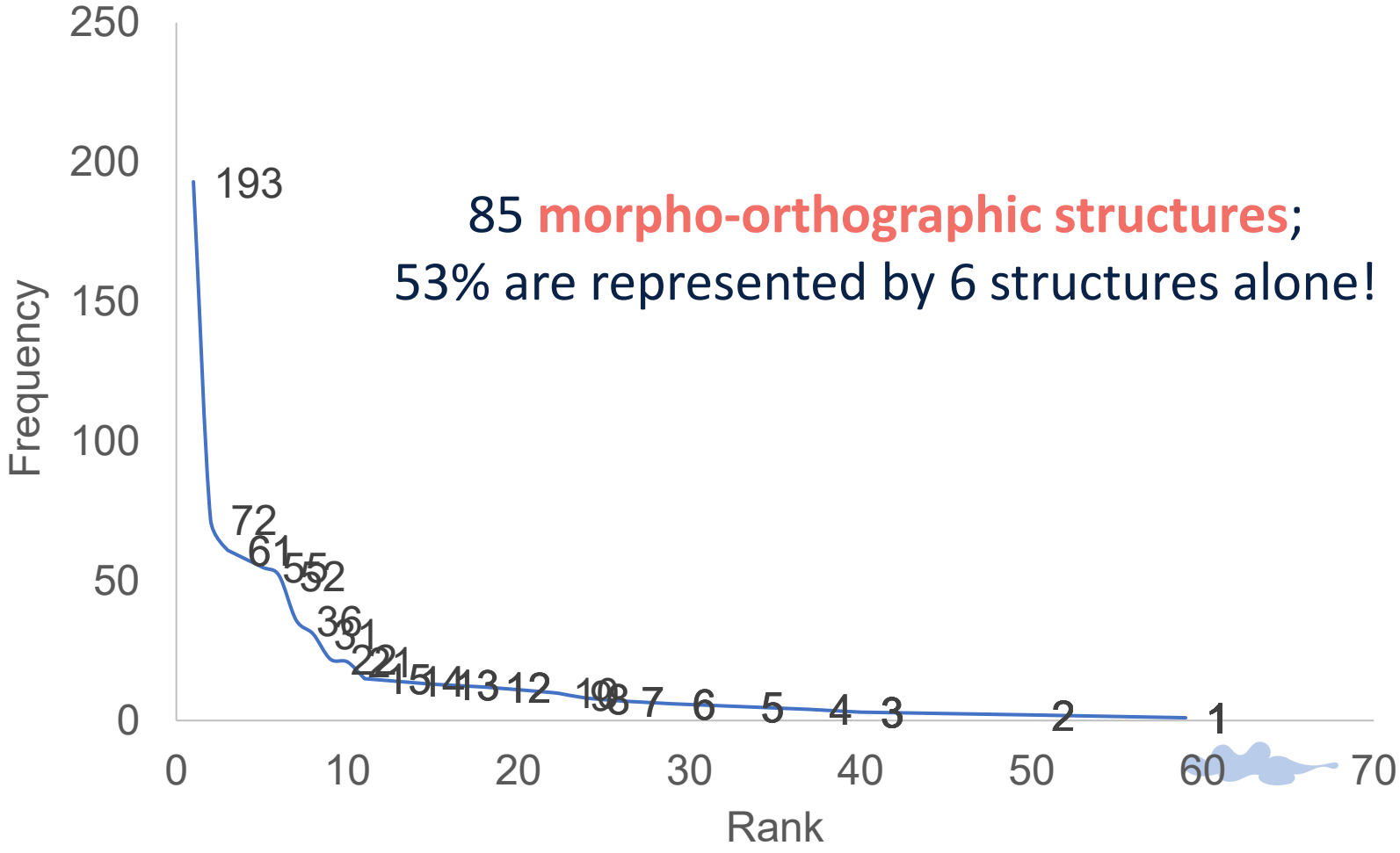
Evanhaim & Bar-On, in preparation

Morpho-Orthographic Density

944 noun types

136 morphological patterns

85 morpho-orthographic structures

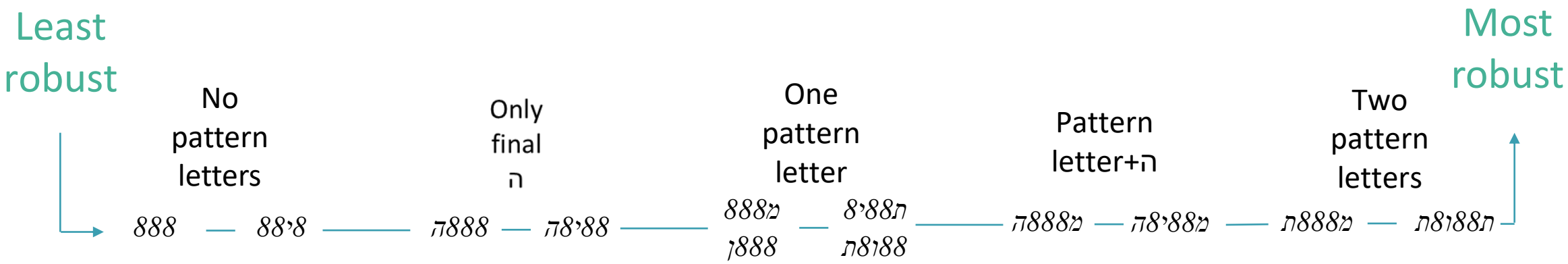


85 **morpho-orthographic structures**;
53% are represented by 6 structures alone!

Evanhaim & Bar-On, in preparation

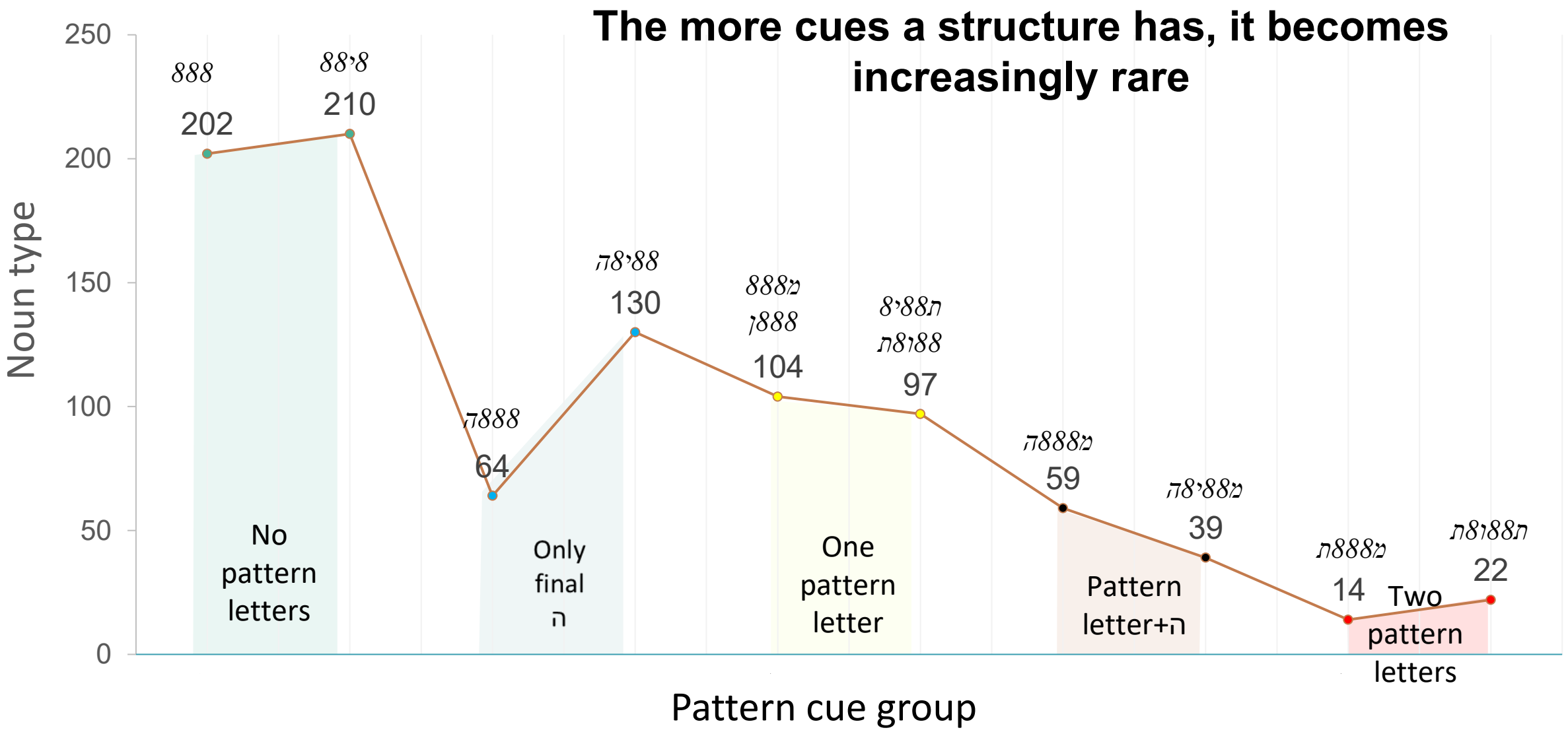
Transparency-Frequency Trade-offs in Hebrew orthography

?How do pattern cues distribute



Evanhaim & Bar-On, in preparation

How do pattern cues distribute?



Evanhaim & Bar-On, in preparation

What do Hebrew readers encounter in the first years of acquisition?

📊 Orthographic Convergence

Thousands of items → Few recurring structures; Zipfian distribution

⚡ Statistical Learning

High-frequency "super-patterns" Enable rapid acquisition

⚠️ The Cost

Frequent = Ambiguous
Homographic competition

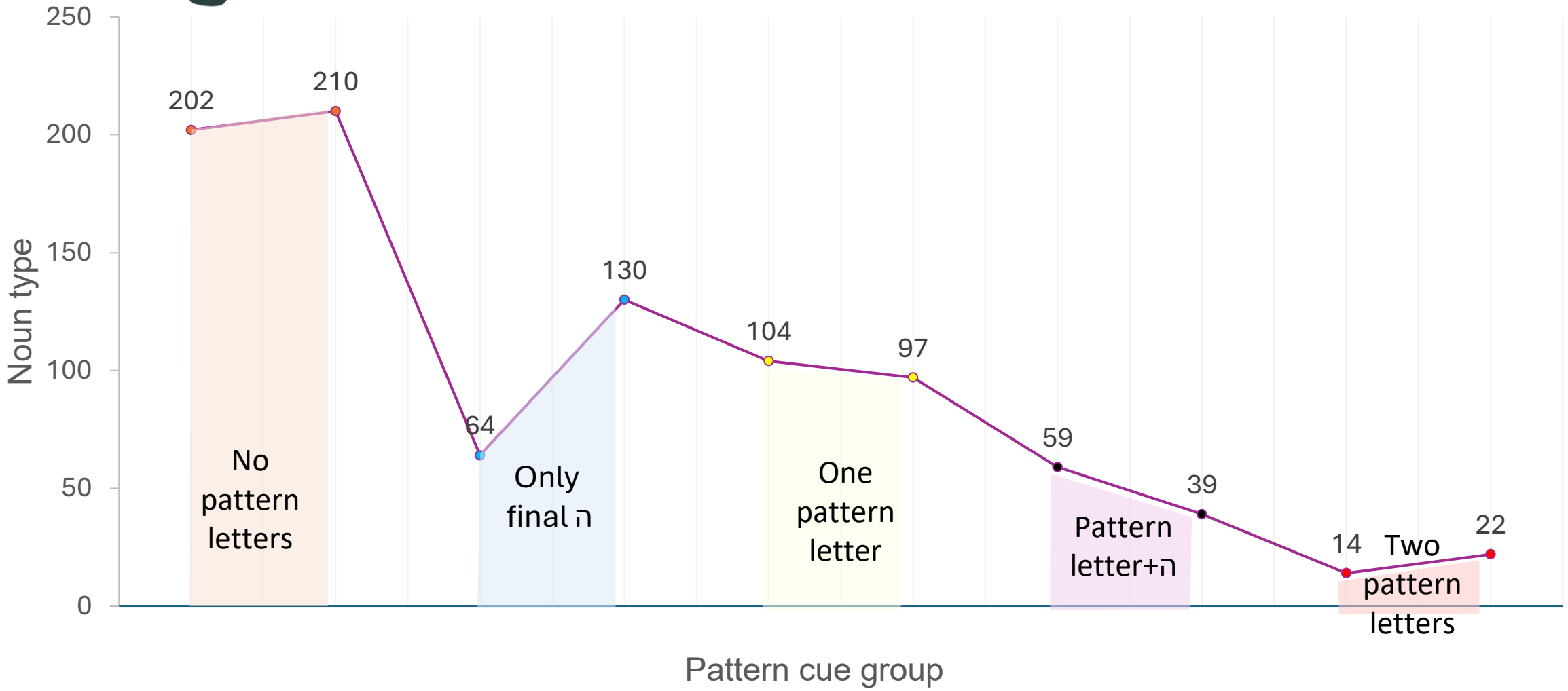
🔍 Low-Frequency Words

Pattern cues become more critical

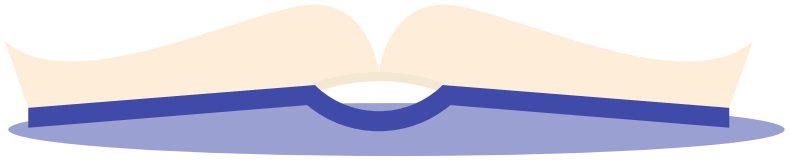
⚖️ The Trade-off

Early fluency via “super-patterns” requires rapid development of disambiguation strategies (morphology, lexicon, context)

Do these different types of morpho-orthographic forms affect reading performance differently?



Reading Experiment



Pseudo-Nouns Decoding

✓Lexical “bypass” – enhanced inspection of the morphological contribution

✓Multiple vowel pattern interpretations are possible

STIMULI

Morphological-based Pseudo-nouns

(long)	התקלגות	דרקלגוס
(medium)	מדלר	גדלר
(short)	קומס	תוקס

STIMULI

Reading Experiment Stimuli

pattern-letters at both word boundaries (Long)	pattern-letters at one word boundary (Medium)	Three vs. two-root-letter structures (Short)
10 morphologically structured	10 morphologically structured	10 canonical
10 non-morphological controls	10 non-morphological controls	10 allomorphic

pseudo-nouns varying in morphological complexity and opacity

60 pseudo-nouns were embedded in 30 short paragraphs

STIMULI

אם תמצאו חיה פצועה קחו אותה אל **המסוד** שבמרכז העיר במהרה.

שם היא תקבל **מקלוגת** שתעזור לה לחזור לטבע.

If you find an injured animal, take it to the **MSWD** downtown quickly.

There it will receive a **MKLWGT** that will help her return to nature.

STIMULI

אם תמצאו חיה פצועה קחו אותה אל **המסוד** שבמרכז העיר במהרה.

שם היא תקבל **מקלוגת** שתעזור לה לחזור לטבע.

If you find an injured animal, take it to the **MSWD** downtown quickly.

There it will receive a **MKLWGT** that will help her return to nature.

POPULATION

- ✓ Originally built for children
- ✓ For a reference point and validation – was tested on adults

- ➢ 40 adult participants (20 women; age mean 29 years)
- ➢ Native Hebrew (no language, Hearing or learning disabilities)

Adult and children participants read the sentences using the ReadLet platform

PROCEDURE

ReadLet

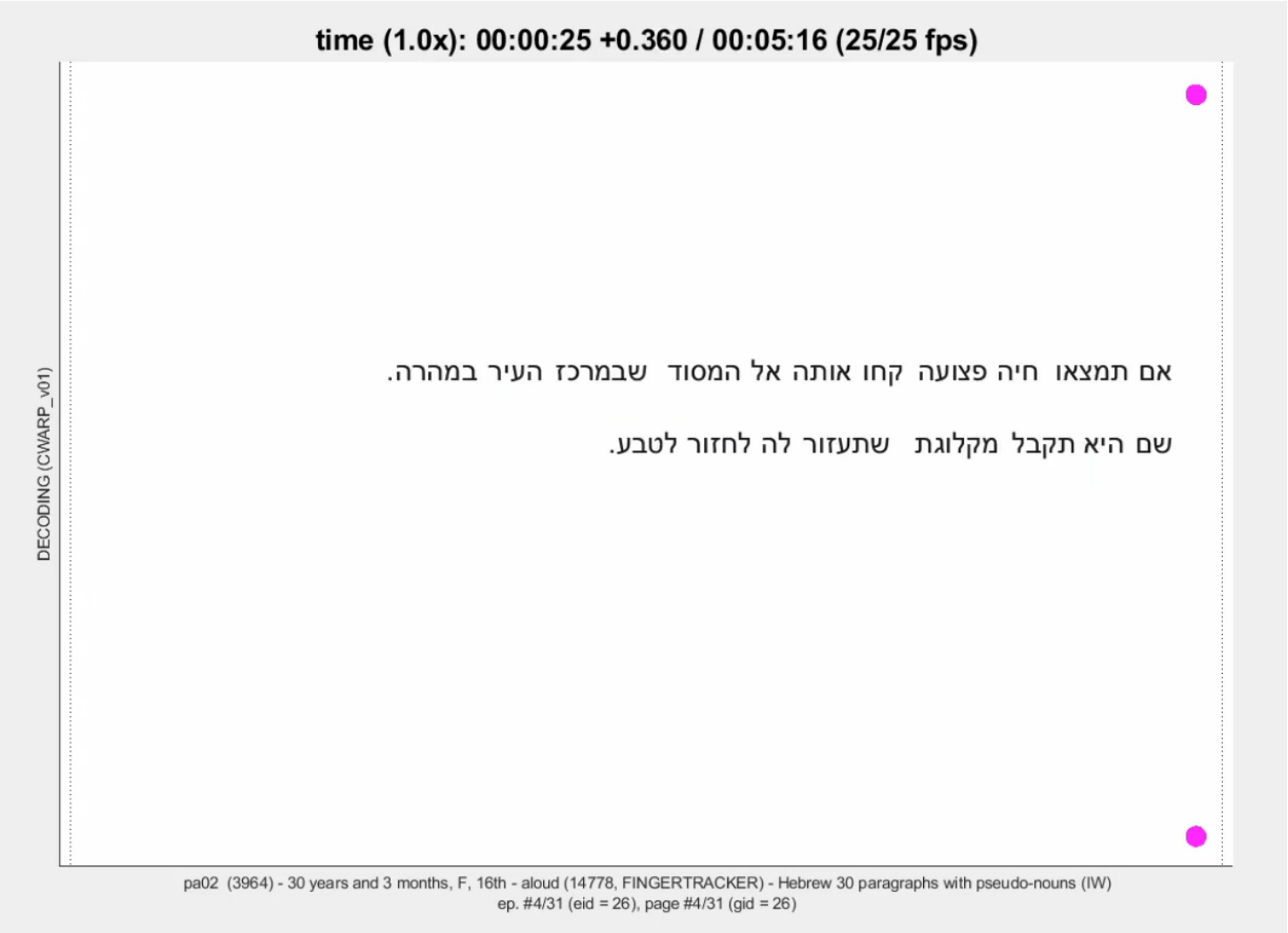
An innovative platform for accurate, evidence-based reading assessment

Vito Pirrelli, Claudia Marzi, Marcello Ferro, Claudia Cappa, Loukia Taxitari, Andrea Nadalini, Ester Bruno, Davide Crepaldi, Viola del Pinto

Ferro et al., 2018 ; Crepaldi et al., 2021

PROCEDURE

ReadLet



- ✓ Voice recording
- ✓ Finger sliding time
- ✓ Time of reading
- ✓ Speech-to-text algorithm

- ✓ Highly correlates with eye tracking
- ✓ Highly ecological, non-invasive protocol for extensive data elicitation

Ferro et al., 2018 ; Crepaldi et al., 2021

ReadLet Finger-Tracking Method

First Application in Hebrew

👶 Child-Friendly

Intuitive finger-tracking replaces artificial paradigms

🎯 Ecological Reading

Single-word RT within natural sentence context

🔍 Morphological Precision

Separates definite article (ה) reading time from noun reading time

Originally developed for Italian | Adapted & validated for Hebrew

✓ Validated for Hebrew

✓ Length effects

✓ Frequency effects

✓ Adults & children

✓ Voice-finger correlation (Bland-Altman)

PROCEDURE

ReadLet → Word reading speed

Decoding possibilities transcription → Decoding diversity

Findings

Adult Reading

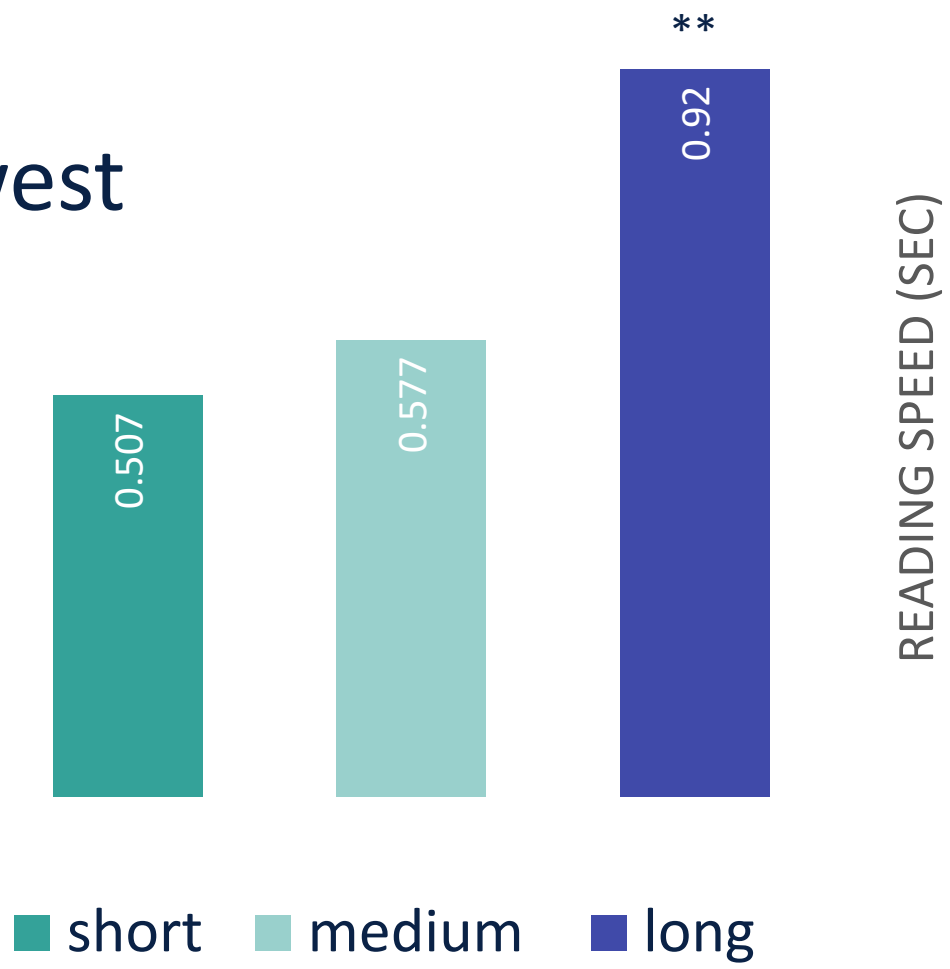
READING SPEED

Length Effect

Long words were read the slowest

קומס = מדלר < התקלגות

תוקס = גדלר < דרקלגוס



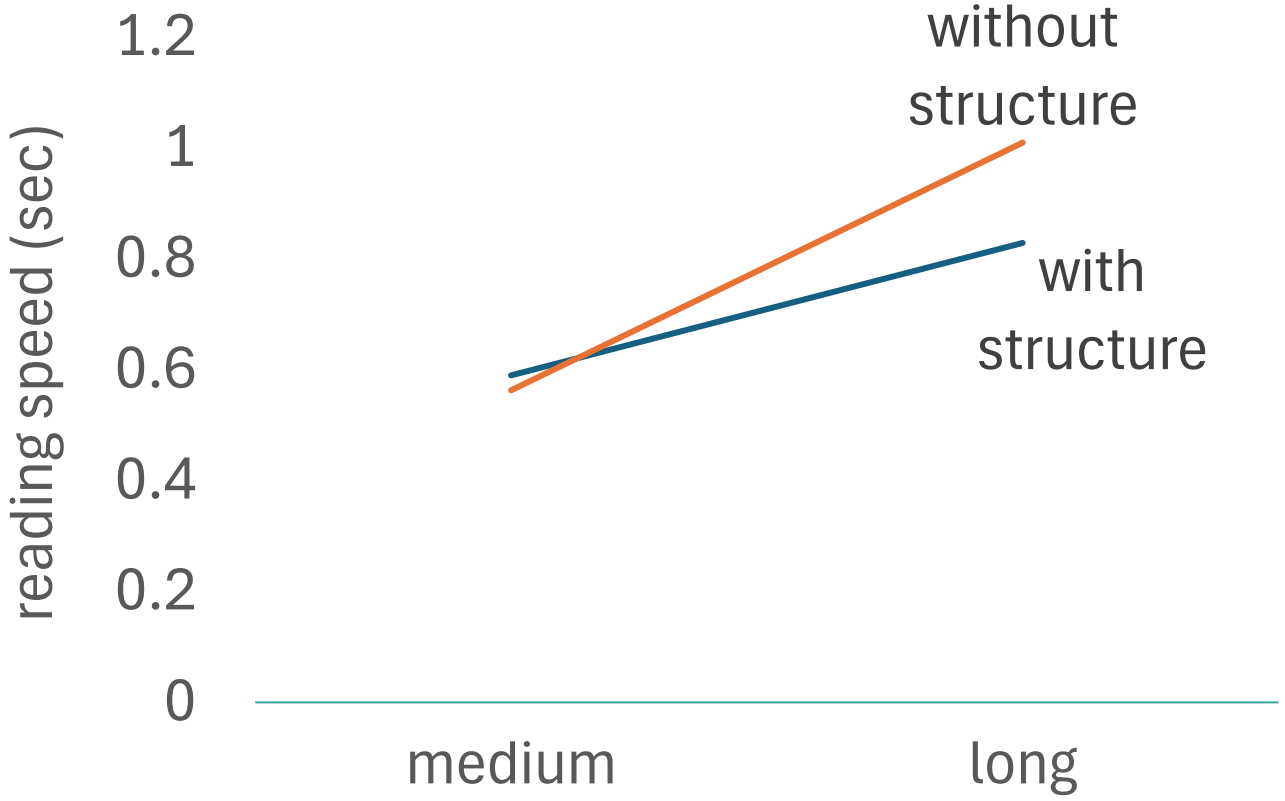
READING SPEED

Morphological Effect

Morphological structure facilitated faster reading only with pattern cues at both word boundaries

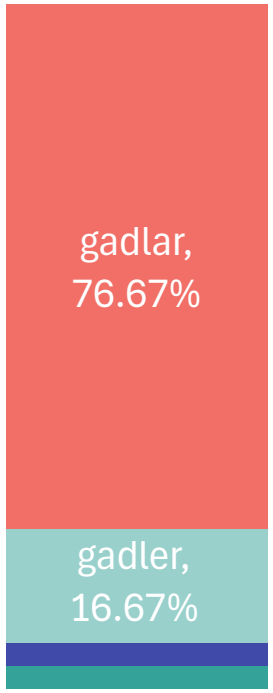
דרקלגוס > התקלגות

גדלר = מדלר

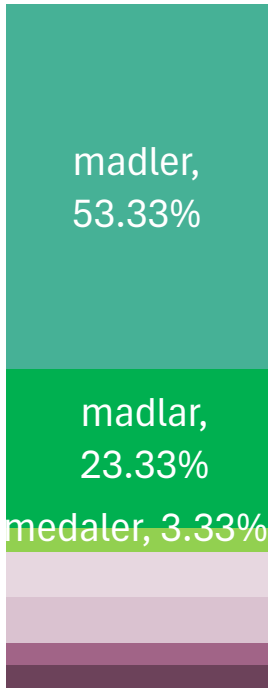


Decoding diversity

גדלר vs מדלר

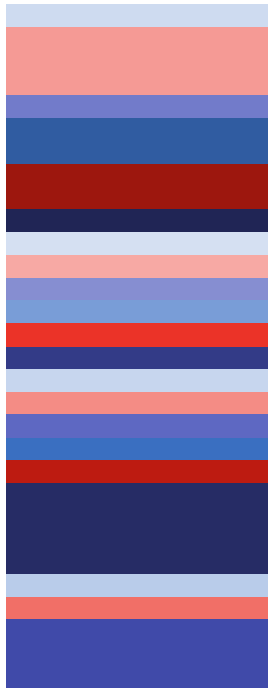


4 possibilities



7 possibilities

דרקלגוס vs התקלגות



21 possibilities

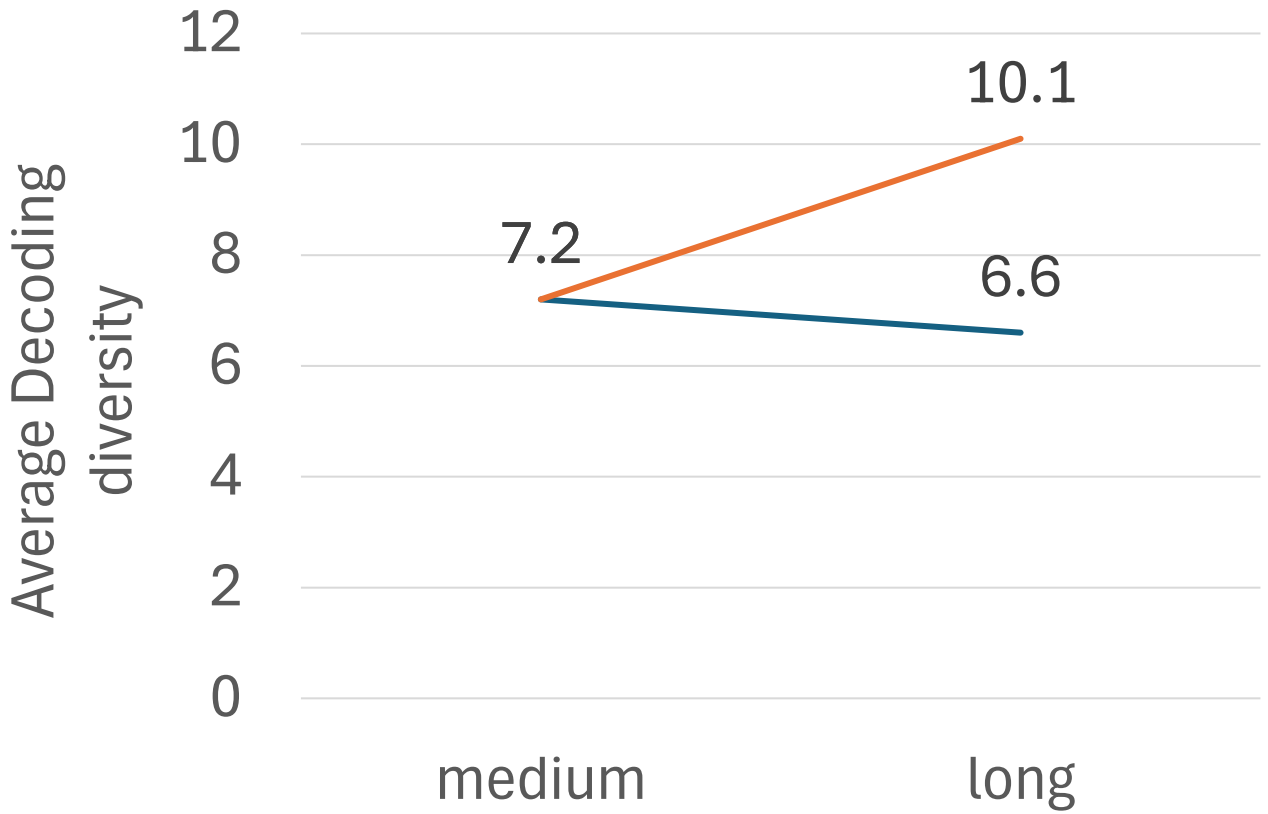
2 possibilities

Decoding diversity

Morphological structure contributes to decreased decoding diversity only with pattern cues at both word boundaries

דרקלגוס < התקלגות

גדלר = מדלר



Findings summary

Expected		Unexpected
Length Effect	→	No difference between medium and short
Morphological Effect	→	Only in long words (full patten envelop) Reading speed + decoding diversity

Orthographic Density

2 sources for structure repetitiveness

Morphological structures

Semitic + Universal

(In our study contributed only for long words)

מברג *maCCeC*

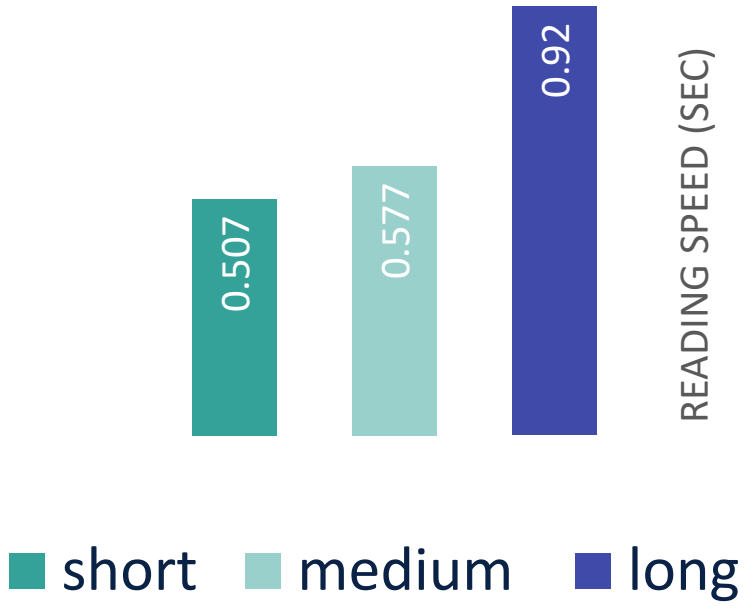
Phonological structures

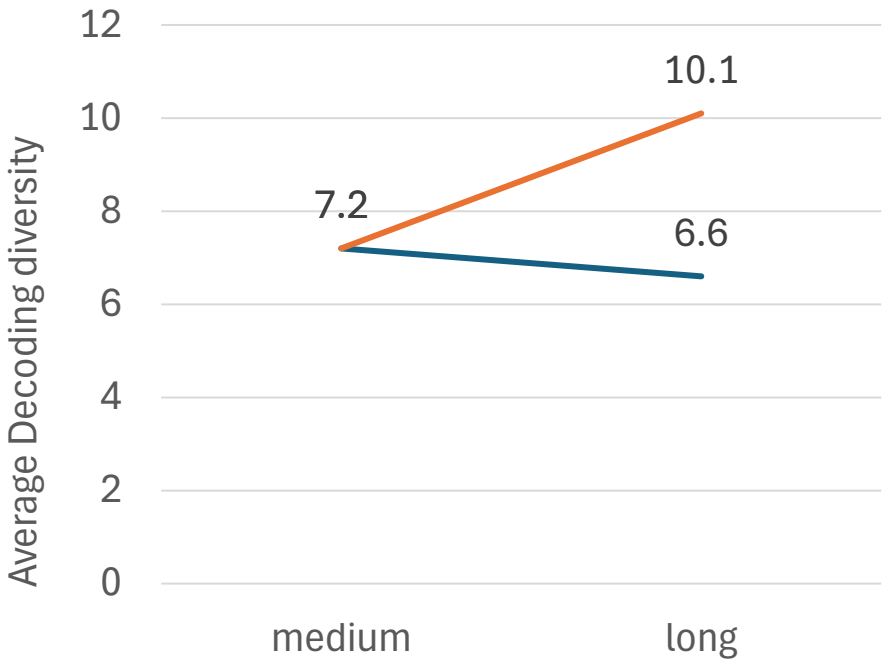
Two-syllabic structure

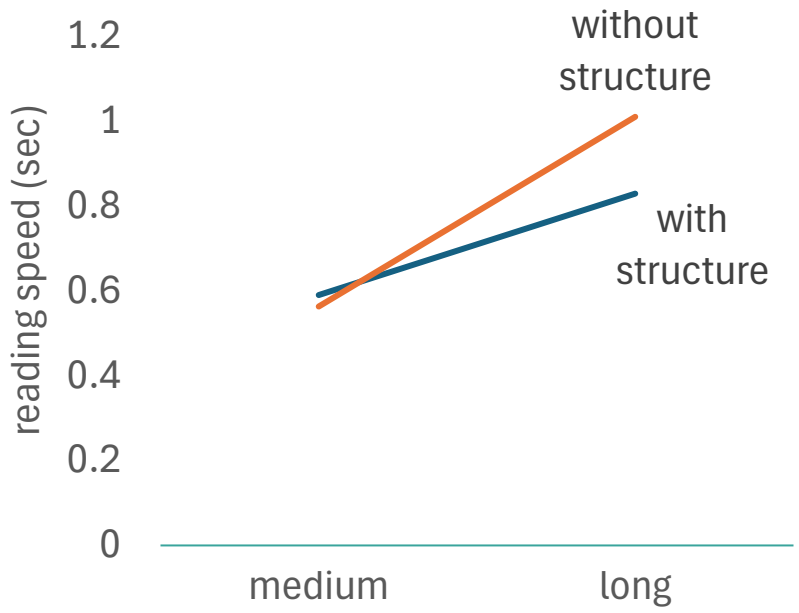
888מ / 8888

CVC.CVC

ברזל *barzel*


READING SPEED (SEC)
0.507
0.577
0.92
short
medium
long


Average Decoding diversity
12
10
8
6
4
2
0
10.1
7.2
6.6
medium
long


reading speed (sec)
1.2
1
0.8
0.6
0.4
0.2
0
without structure
with structure
medium
long

What's New?

🔍 For Hebrew

Our study **qualifies** morphology's central role - spotlighting **recurring phonological structures** previously overlooked. These overlap with noun patterns, **expanding Hebrew's structural repetitiveness**, strengthening "super-patterns," and accelerating orthographic learning

Disclaimer - pseudo words

Morphological pattern (structure) +

pseudo root (semantic)

What's New?

🌍 Universal Implications

How do writing systems serve readers? We defined **orthographic convergence**: readers acquire a constrained set of structures converging to recurring patterns. **How universal is this?**

*Hebrew's five-vowel system naturally limits diversity

Thank You for listening!

naamaeh@mail.tau.ac.il