ROOT FOR A CHANGE

Teaching root awareness to Hebrew L1 young children with interactive videos

NOFAR RIMON[1], IRIS MELAMED GORODESKY[2], GOREN GORDON[3,4] & EINAT GONEN[2]

[1]Department of Linguistics, Harvard University, Cambridge, MA, USA
[2]Department of Hebrew Language, Tel Aviv University, Tel Aviv, Israel
[3]Department of Informatics, Luddy School, Indiana University, Bloomington, IN, USA
[4]Curiosity Lab, Department of Industrial Engineering, Tel Aviv University, Tel Aviv, Israel

Abstract
The Semitic root forms an integral part of Hebrew language instruction, and though formally taught from second grade on, its extraction seems to be an obstacle for most students. We used our unique online root awareness instruction program in order to teach young children to extract the roots of different Hebrew words, as well as pseudowords, and to detect their weak spots in the process. In this paper, we present the effects of our program, as demonstrated by a significant improvement in the children's root extraction skills. We further provide an analysis of the children's responses, based on which we specify the factors that best predict failure in root extraction. These factors may lead to new insights regarding Hebrew language instruction as well as other Semitic languages.

Keywords: Language education, L1 instruction, morphological awareness, Hebrew morphology, educational intervention

Rimon, N., Melamed Gorodesky, I., Gordon, G., & Gonen, E. (2026). Root for a change: Teaching root awareness to Hebrew L1 young children with interactive videos. L1-Educational Studies in Language and Literature, 26, 1-26. https://doi.org/10.21248/l1esll.2026.26.1.1034
Corresponding author: Nofar Rimon, Department of Linguistics, Harvard University, Cambridge, MA, USA. Email: nofarimon@gmail.com

1. INTRODUCTION

Traditionally, one of the main characteristics of Semitic languages that differentiates them from other language families has been their non-concatenative morphology (Berent & Shimron, 1997; McCarthy, 1981). As a Semitic language, Hebrew is known for its root-and-pattern morphology, meaning that words are composed of three or four consonants, referred to as *root*, combined in a vocalic pattern referred to as *binyan* for verbs, and *mishkal* for nouns, often with additional affixes. While Modern Hebrew has only seven verbal patterns that inflect in three tenses—past, future and present—there are hundreds of nominal patterns.

Morphology is one of the organizing principles of the mental lexicon (Marslen-Wilson, 2016; Ravid & Schiff, 2006), and research shows that in the mental lexicon of Hebrew and Arabic speakers, the root forms a separate entity that is essential for reading and writing (Berent & Shimron, 1997; Bolozky, 1999; Frost et al., 1997; Mimouni et al., 1998; Ravid, 2003; Ravid & Schiff, 2006). The notion of the root is taught in schools from second grade on (Lipkin, 1985; Ravid, 2003), and students are required to extract the roots of different verbs and nouns in language examinations. Therefore, students must have explicit metalinguistic knowledge of the root, although root perception and awareness are inherent to native speakers of Hebrew (Ravid & Schiff, 2006). We treat morphological awareness as a subcategory of metalinguistic awareness, referring to learners' ability to analyze and manipulate morphemes—a skill that supports language learning (Akbulut, 2019; Amirjalili & Jabbari, 2018; Carlisle, 1995; Derwing et al., 1995; Yucel-Koc, 2015).

Several studies show that children as young as three can interpret and coin novel words by relying on the root (Bentin & Frost, 2013; Berman, 1999, 2003; Clark & Berman, 1984). However, the pre-test results of a previous study (Levinson et al., 2021, 2025) demonstrate that children's ability to explicitly state the root of a word is very limited, with an average of correct answers of 1.98 (out of 10). Surprisingly, this applies not only to 5–6-year-old preschools, who had not yet been formally introduced to roots, but also to 7–8-year-old second graders, who had already received such instruction and had mastered basic reading and writing.

Intervention studies found that explicit morphological instruction positively influenced children's morphological awareness and literacy skills in Hebrew and Arabic (Levin et al., 1999, 2001, 2008; Ravid & Malenky, 2001; Schiff & Saiegh-Haddad, 2018; Vaknin-Nusbaum & Raveh, 2019; Vaknin-Nusbaum, 2021), and in other non-Semitic languages (Bowers et al., 2010; Carlisle, 2010; Kirk & Gillon, 2009). Since existing school methods do not appear to bring children to the required level of root extraction skills, which is dominating root extraction by the end of second grade, and due to the relevance of the root to morphological awareness, literacy skills and language curriculum, we created a unique root instruction program. The program teaches children of a young age to extract roots from verbs, nouns, and gerunds of different morphological patterns in Hebrew, in only five short lessons.

Teacher training is not required, as the program is fully accessible online as an extracurricular activity yet could be easily implemented in schools.

Preliminary findings from a previous study conducted in a summer camp with social robots indicated the success of our program (Levinson et al., 2021, 2025), as children scored a significantly higher score in the post-test. There, physical social robots, robots whose purpose is to interact with people, and which were used in educational settings, instructed children aged 6-8 years in a summer camp activity. Similar preliminary results were obtained in another study where our program was adapted to teach root extraction to Arabic speaking children aged 6-8 years, also using social robots (Abu-Mukh et al., 2023; Gonen et al., 2025). In the current intervention study, we strived to address the limitations of the previous Hebrew study and to test our program in a manner that would allow us to reach a larger and more varied audience. Therefore, the experiment was conducted completely online, using various E-learning technologies instead of social robots. The current experiment was much more comprehensive than the previous Hebrew experiment, comprising a larger variety of experimental items—34 instead of 18. This also enabled a more fine-grained examination of the factors and the challenges to root extraction as in this study children dealt with more complex forms such as pseudowords, quadriliteral roots and weak roots. Unlike school pedagogy, which focuses on identifying roots in written words, our program is entirely oral and requires no reading or writing skills.

The effects of our program on children's active morphological awareness were tested by evaluating their knowledge prior to and after the program. All children were L1 Hebrew speakers, and some had already received explicit root instruction. We examined the children's metalinguistic ability to extract the roots of Hebrew words that differ from one another in their morphology, lexical category, and semantic meaning, inter alia. We then sought the factors that made root extraction more difficult using the children's responses. These factors have theoretical implications for linguistic research and practical implications for education—understanding what properties make a word more challenging for root extraction and improving morphology instruction in the future. In this paper, we present the effects of our program, as demonstrated by the significant improvement in children's root extraction abilities, the analysis of the children's responses and the predictors of failure in root extraction. Initial outcomes of the experiment have been presented in Rimon et al. (2022) yet will be analyzed and discussed in detail in this paper.

The paper is structured as follows: Section 1 sets the stage, providing background about Hebrew morphology and root instruction in schools, and presenting our research questions and hypotheses. Section 2 presents the methodology used in the experiment. Section 3 provides the results. In section 4 we connect all the dots and provide a general discussion. Section 5 concludes.

1.1 The Hebrew root

Hebrew is traditionally known for its non-concatenative morphology, meaning that words are composed of a root, three or four consonants that are integrated in a vocalic pattern. This is opposed to the concatenative morphology of languages like English, where morphemes are concatenated linearly. Semantically, the root carries the basic meaning of the word, while the pattern classifies the word into a specific category, as do English derivational suffixes (Berman, 1993; Deutsch & Kuperman, 2019; Levie et al., 2019; Ravid & Schiff, 2006). There are seven verbal patterns, also known as *binyanim* (sg. *binyan*) and many nominal patterns, also known as *mishkalim* (sg. *mishkal*).

Table 1. The root k-t-b in various patterns (exhibiting k-χ and b-v alternations). C represents a placeholder for a root consonant

Pattern	Pattern Type	Word	Meaning
CaCaC	verbal	*katav*	wrote
CiC(C)eC	verbal	*kitev*	added someone as a recipient
hiCCiC	verbal	*hiχtiv*	Dictated
niCCaC	verbal	*niχtav*	was written
hitCaC(C)eC	verbal	*hitkatev*	corresponded (with someone), exchanged letters
CuC(C)aC	verbal	*kutav*	was added as a recipient
huCCaC	verbal	*huχtav*	was dictated
CCaC	nominal	*ktav*	Script
CaC(C)aCa	nominal	*katava*	Article
CCiC	nominal	*ktiv*	Spelling

Roots are usually composed of three consonants that appear throughout the inflection of the verb or the noun, yet there are also weak roots from which approximately 10% of Hebrew words are derived, many of which are high frequency words (Frost et al., 2000; Levie et al., 2019; Schiff et al., 2008). In weak roots, one radical fails to surface in some forms within the paradigm. These radicals are usually glides, other sonorants or pharyngeal consonants. For example, the verbs *nafal* “he fell” and *jipol* “he will fall” share the root *n-f-l*, yet the first root radical surfaces in the former but not in the latter. Weak roots are the major challenge in morphology instruction—they are not intuitive for native speakers and must be taught explicitly (Ravid, 2003), and they are substantially harder to retrieve from the lexicon (Ravid & Bar-on, 2001). Another crucial challenge is distinguishing root consonants from pattern or affix consonants. Root consonants can appear in any position within a word, while pattern or affix consonants typically occur at the beginning or end, and the same consonants may function as part of a root morpheme in one word and as a pattern or affix morpheme in another (Deutsch et al., 2018).

The root has a psycholinguistic reality for native speakers of Semitic languages, especially for preliterate preschoolers that have not received a formal instruction of this subject (Berent & Shimron, 1997; Berman, 1993, 1999, 2003; Boudelaa, 2014; Clark & Berman, 1984; Faust, 2019; Frost et al., 1997; Ravid & Malenky, 2001; Taha & Saiegh-Haddad, 2017). Research has shown that children rely on the consonantal skeleton while interpreting and producing unfamiliar words and creating associations between lexical entries (Bentin & Frost, 2013; Berman, 1999, 2003; Clark & Berman, 1984).

Nonetheless, previous studies show that children’s active root extraction skills are very limited, indicating the shortcoming of root instruction in schools both in L1 Hebrew (Levinson et al., 2021, 2025) and in L1 Arabic (Abu-Mukh et al., 2023; Gonen et al., 2025). As stated above, intervention studies found that explicit L1 morphological instruction positively influenced children's morphological awareness and literacy skills in these languages. Therefore, reinforcing children's active root extraction abilities would help them achieve a deeper understanding of morphology and facilitate word acquisition and processing (Ravid et al., 1999, 2016), reading comprehension (Ravid & Schiff, 2006; Vaknin-Nusbaum, 2021) and spelling (Levin et al., 2001).

1.2 Root instruction in schools

Roots and word families are formally taught in schools in Israel starting from second grade, thus familiar to native speakers of Hebrew and Palestinian Arabic (Bolozky, 1999, 2003; Ravid, 2003; Ravid & Malenky, 2001). At first, root instruction focuses on regular triconsonantal roots and identifying semantic relations between words. It is only in grades 10 to 11 that students encounter weak roots and learn different strategies to deal with them. Generally, quadriliteral roots receive much less emphasis than triconsonantal roots, and nouns receive much less emphasis than verbs (Rosner, 2003).

One basic root extraction method taught in schools relies merely on orthography—children are taught to erase the letters that do not belong to the root or to mark those that do. This method is very limited as it leads to a dependency on orthography, which distances native speakers from the perception of both the root and the pattern as intuitive independent entities.

A 2023 report by the Central Bureau of Statistics on subject-teacher compatibility in Hebrew, mathematics and English indicates that in 2022 only 19.9% of Hebrew teachers in secular elementary schools were adequately trained for Hebrew instruction, compared with just 7.8% in religious schools (Central Bureau of Statistics, 2023). Teacher shortages further increase reliance on underqualified or out-of-field staff, thereby compromising workforce quality (Darling-Hammond et al., 2023; Donitsa-Schmidt, 2025; Flores & Craig, 2023). In light of these conditions, together with the shortcomings of Hebrew and Arabic root instruction in schools indicated by previous studies (Abu-Mukh et al., 2023; Levinson et al., 2021, 2025; Gonen et al., 2025), we developed an extracurricular program that does not rely on qualified language teachers. Importantly, the program is not intended to replace classroom teachers but to serve as a complementary instructional tool that supports explicit metalinguistic instruction. Accordingly, the present study does not aim to compare online and frontal instruction.

Our new teaching method is based on oral instruction rather than orthography and allows root instruction to begin as early as preschool. Choosing an oral method was meant to provide the children with unmediated access to the root. This access may help with reading and writing acquisition, as shown in Vaknin-Nusbaum et al. (2016), yet the effect of our program on these skills was not tested. In order to scrutinize the efficiency of our method, we conducted an extensive intervention study that tested the children's ability to orally extract the roots of diverse forms of different levels of complexity.

1.3 Research questions and hypotheses

Given the insufficient outcomes of root instruction in schools demonstrated by low achievements in a previous pre-test (an average number of correct answers of 1.98) (Levinson et al., 2021, 2025) and considering the importance of the root to morphological awareness, literacy skills, the mental lexicon and school curricula, our intervention study aimed to answer the following questions:

1) Does our online root instruction program significantly improve children's ability to extract roots of words of different levels of difficulty (e.g., nouns, pseudowords and weak roots) in Hebrew?
2) What are the linguistic factors (e.g., word frequency, pattern, lexical category, etc.) that influence children's success in root extraction?

We predict, based on a previous study (Levinson et al., 2021, 2025), that after participating in our program children will significantly improve their root extraction abilities and will be able to successfully deal with more complex forms as well, regardless of age group. As for the linguistic factors that contribute to difficulty, we predict nouns to be more challenging than verbs, since verbs are commonly manipulated (conjugated). We further predict weak roots and prefixes to pose challenges. As previously mentioned, weak roots and prefixes are a primary obstacle in root instruction. We hypothesize that children would struggle with restoring the missing root radical in weak roots and misinterpret a prefix as a root radical.

To address the first research question, we will conduct a pre-test and a post-test. A significant difference between the tests will indicate an improvement in the children's root extraction skills following our program. Including the pre-test items in the post- test provides a more direct basis for comparison and assessment of their success. Further, items of a higher level of difficulty will be compared to their counterpart low-difficulty items (i.e., pseudowords and weak roots will be compared to their counterpart regular forms), and the lack of a significant difference between them would point at success in dealing with different levels of complexity in root extraction. To address the second research question, we will analyze the children's responses and run a logistic regression that would reveal the factors that contributed to failure. These factors will be based on the different variables that distinguish our experimental items from one another. Identifying what makes root extraction difficult can help educators design more effective curricula.

2. METHODOLOGY

2.1 Technology

We prepared five interactive videos using recorded audio text and avatars from SitePal.com©, a commercial website for integration of audio files and animated avatars. We used two cartoon-like avatars to address the appropriate age range. The recorded videos were uploaded to YouTube.

2.2 The lessons

The program consisted of five video lessons to cover all the topics that we wished to examine. Each video was six to ten minutes long, and the overall duration of all lessons was approximately 40 minutes. We kept the lessons relatively short to maintain the children's interest and engagement. The lessons were divided by topics in an increasing level of difficulty, based on previous findings (Levinson et al., 2021, 2025). The first lesson focused on the patterns *CaCaC* and *CiC(C)eC*, the second one focused on *hitCaC(C)eC* and *niCCaC* and the third on *hiCCil*, *CuC(C)aC* and *huCCaC*. The fourth lesson taught pseudowords and nominals and the final lesson focused on quadriliteral roots, more nominals and a summary.

In the videos, the children encounter a long series of structured examples of words and their roots (see Appendix III). The avatar poses the question "what is the root of the word X?", then, after giving the children a few seconds to answer, the avatar provides the correct answer, adding a compliment after each set. Crucially, the avatars could not hear the children's responses. The avatar also tells the children jokes and asks them to do some activities during the lesson to keep them engaged, like asking them to stand up or jump, and then to provide the roots of these activity verbs. The questions and answers are given in a rhythm that resembles recitation, and the cartoon avatars encourage children's active participation, as they decipher the roots alongside the avatars. Thus, children learned to extract roots based on oral instruction, theoretical explanations were not provided and reading or writing skills were not required.

The children were instructed to watch the videos and encouraged to answer by the avatar. The parents were instructed to play the videos for the kids and watch them while they learn without intervening in their learning process. After each lesson, the parents had to fill in a feedback form to get the link for the following lesson. In the form, the parents had to report how the lesson went, how concentrated the child was, whether the child actively participated in the class (i.e., answered the avatar's questions out loud, followed the avatar's instructions), add comments and report any unusual occurrence throughout the lesson.

2.3 Participants

Ninety-six children took the pre-test and completed the first lesson, 84 completed the second, 73 completed the third, 69 completed the fourth, and 68 completed the fifth. The final analysis was performed on 64 children (32 males and 32 females) who completed both the pre-test, the post-test and all lessons. The children were preschoolers (n = 29), first graders (n = 23) and second graders (n = 12), aged 5;3-8;9 (M = 6;11) and came from different regions and varied backgrounds. The preschoolers had no reading or writing skills, whereas the second graders had already mastered both and had been introduced to the notion of roots in school. We chose this age group with the intention to examine whether the younger children could be challenged with a higher level of linguistic education and draw level with the second graders. All participants were native speakers of Hebrew who live and grew up in Israel: 55 monolinguals and 9 bilinguals with dominant Hebrew, with no known language impairment. The experiment was run during summer break, when children do not have any school classes and could not learn the material taught in the study elsewhere.

The study was approved by the Institutional Review Board (IRB) of Tel Aviv University.

2.4 Procedure

An invitation to participate in the experiment was posted on social media and sent to potential participants individually. Parents had to consent to participate and submit a form of demographic and personal information. First, the children had a one-on-one Zoom meeting with an interviewer, where they watched a short introductory video inviting them to join the 'Root Detective' club and received a short training of three examples of root extraction. Then, the children were asked to extract the roots of 13 words, without additional context for the word. All questions were of the same form—"What is the root of the word X?". The interviewers always gave positive feedback to the children, regardless of their answers. After the pre-test, the children participated in five root instruction lessons in their free time on YouTube. After each lesson, the parents submitted a feedback form to get the link for the following lesson. Later, having completed all five lessons, the children were invited to another one-on-one Zoom meeting, where, after a short video they took the post-test. In the post-test, the children had to extract the roots of 34 words. All interviews were recorded and reviewed by the interviewers to make sure that the children's answers were correctly transcribed and to ensure inter-rater reliability.

2.5 Data analysis

The children's responses were recorded and annotated by the interviewers. Answers were classified as either correct or incorrect (1 or 0), and incorrect answers were later classified by type of error. A full correct answer included all root radicals in the right order, in either *CaC(C)aC* or *CiC(C)eC*. Partial responses were not counted as correct. Since our program is oral, the children were taught to provide the roots in the pattern *CaCaC*, to facilitate pronunciation and create a uniform rhythm to the answers that children can rely on. Additionally, stop-fricative alternations of the stops /b/, /k/ and /p/, which alternate with [v], [χ] and [f] respectively, were not taken into account, since Modern Hebrew exhibits some unstable alternations (see Albert, 2019, Berman & Bolozky 1978, p. 2, and Temkin Martinez, 2010, for further discussion). As long as the child provided one of the alternants, the response was considered correct.

2.6 Materials

2.6.1 Pre-test

The materials for the pre-test comprised 13 items, having 11 verbs and 2 nouns. There was a verb of a triliteral root for each of the seven verbal patterns, two verbs of a quadriliteral root in *CiC(C)eC* and two additional pseudo-verbs in *CiC(C)eC* and *hiCCiC*. The two nouns were of the patterns *taCCiC* and *maCCeCet*. All pseudowords used in the experiment were composed of an existing pattern and a non-existent root, also referred to as pseudo-root. Except for the two verbs whose roots were quadriliteral, all other words were of triliteral strong roots (i.e., all their radicals appear throughout the paradigm). The verbs were either inflected in the past, present and future or uninflected.

2.6.2 Post-test

The materials for the post-test comprised 34 items: 24 verbs, 8 nouns and 2 gerunds. All pre-test materials were included in the post-test. Seven of the verbs were pseudo-verbs (two of which appeared in the pre-test), so that each verbal pattern had a pseudo form, and two of the nouns were pseudo-nouns. Five verbal forms were in

CaCaC: one pseudo-verb, one regular verb of a triliteral root, and three weak-root forms of three different root classes. All forms were transparent, i.e., the missing root radical could be restored from the past tense form. Another five verbal forms were in *CiC(C)eC*: one pseudo-verb, two regular verbs of a triliteral root, and two verbs of a quadriliteral root. *hitCaC(C)eC*, *hiCCiC*, *huCCaC* and *niCCaC* had three forms each: two regular verbs of a triliteral root and one pseudo-verb. Finally, two verbal forms were in *CuC(C)aC*: one pseudo and one regular. As for the nouns, we chose four nominal patterns—two with affixes and two without, in order to examine the effect of affixation on root extraction from nouns. The two gerunds were in *hiCCiC* and *hitCaC(C)eC*. Importantly, the materials of the pre- and post-test did not appear in the lessons. Furthermore, the frequency of each item was assessed and controlled for based on data from heTenTen 2021, a 10+ billion token web-crawled Hebrew corpus (Jakubíček et al., 2013).

It should be noted that the pre-test was significantly shorter than the post-test since previous pre-tests indicated extremely low achievements (Abu-Mukh et al., 2023; Gonen et al., 2025; Levinson et al., 2021, 2025). To avoid frustration and resentment towards our program, we shortened the pre-test and extended the post-test, aiming to gather enough data for a more comprehensive analysis of the children's responses at the end of the program.

2.7 Limitations

The study did not have a proper control group. However, because the pre-test and the post-test were administered only two weeks apart, it is unlikely that the participants learned the material elsewhere during that time (summer break), or that developmental changes occurred that could influence the drastic improvements achieved. Further, since the children watched the videos in their free time, we had no unmediated data regarding their concentration and participation, among other things. Nevertheless, we did have the parents' reports of each lesson, in which they included these data according to their impression. Finally, there might have been a socioeconomic status (SES) bias in our study, since we could not control for this factor. Our study was published on social media, and parents who chose to participate in the study, which took place during summer break, are parents who probably view the subject as important and put emphasis on their children's education. Considering this and the fact that the study was conducted online, we believe to have reached families of a middle-high SES. It should also be highlighted that the study was a short-term one, and we did not examine its long-term outcomes.

3. RESULTS

3.1 General analysis

Table 2. Fixed effects estimated from a generalized linear mixed model analyzing success rate by Test, Gender, Grade, and their interaction. Random intercepts for participants and for items were included. Reference levels: Test = Pre, Gender = Male, Grade = Preschool

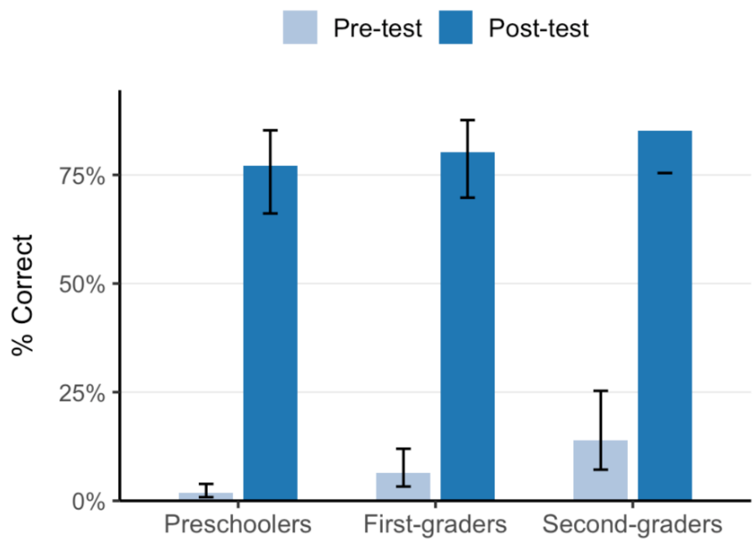
	Estimate	Std. Error	z	*p*
Intercept	-4.348	0.443	-9.81	< 0.001 ***
Test:Post	5.55	0.373	14.879	< 0.001 ***
Gender:Female	0.706	0.295	2.398	0.016 *
Grade:First	1.306	0.361	3.616	< 0.001 ***
Grade:Second	2.172	0.385	5.644	< 0.001 ***
Test:Post:Gender:Female	-0.687	0.293	-2.346	0.019 *
Test:Post:Grade:First	-1.122	0.359	-3.129	0.002 **
Test:Post:Grade:Second	-1.637	0.384	-4.262	< 0.001 ***

A generalized linear mixed model with random intercepts for participants and items fit by maximum likelihood found that the children made a dramatic improvement from the pre-test to the post-test, being 258 times more likely to provide a correct response in the post-test compared to the pre-test (OR = 258, β = 5.55, SE = 0.37, z = 14.88, $p < 0.001$). Mean accuracy increased from M = 0.11 in the pre-test to M = 0.73 in the post-test. A logistic mixed-effects model revealed no difference in post-test accuracy between items that had appeared in both tests and items that appeared only in the post-test (β = –0.14, SE = 0.58, z = –0.24, p = .81, OR = 0.87).

As for demographics, girls performed significantly better than boys in the pre-test, being about twice as likely to provide a correct response (Girls: M = 0.13; Boys: M = 0.08; OR = 2.03, β = 0.71, SE = 0.3, z = 2.4, $p = 0.02$). Post-hoc pairwise comparisons showed that this difference disappeared in the post-test (Girls: M = 0.73; Boys: M = 0.73; OR = 1.02, p = 0.97). The significant interaction indicates that boys made a bigger improvement after the intervention (β = -0.69, SE = 0.29, z = -2.35, $p = 0.02$). Since girls started at a higher baseline, their improvement was smaller, resulting in a reduced gender gap in the post-test.

In the pre-test, first graders performed significantly better than preschoolers (First graders: M = 0.12; Preschoolers: M = 0.04; β = 1.31, SE = 0.36, z = 3.62, $p < 0.001$), and so did second graders (M = 0.23; β = 2.17, SE = 0.38, z = 5.64, $p < 0.001$). This was confirmed by post hoc comparisons: both first graders (OR = 3.69, $p < 0.001$) and second graders (OR = 8.73, $p < 0.001$) outperformed preschoolers. Second graders also outperformed first graders (OR = 2.38, $p = 0.03$). In the post-test, however, no significant differences between the grades were found (Preschoolers: M = 0.7; First graders: M = 0.73; Second graders: M = 0.78; Preschool-First: OR = 0.83, $p = 0.59$; Preschool-Second: OR = 0.59, $p = 0.06$; First-Second: OR = 0.7, $p = 0.32$).

Figure 1: Means of pre- and post-test results by grade



3.2 Linguistic analysis

Table 3. Fixed effects estimated from a generalized linear mixed model analyzing success rate by Prefix and Category. Random intercepts for participants and for items were included. Reference levels: Prefix = Internal, Category = Verb

	Estimate	Std. Error	z	*p*
Intercept	0.866	0.43	2.008	0.045 *
Prefix:Non	1.561	0.688	2.271	0.023 *
Prefix:External	1.27	0.615	2.064	0.039 *
Category:Gerund	-1.085	1.095	-0.99	0.322
Category:Noun	-0.725	0.665	-1.089	0.276

A generalized linear mixed model with random intercepts for participants fit by maximum likelihood revealed that non-prefixed rather than internally prefixed forms were significantly easier to extract roots from (Non-prefixed: M = 0.85; Internally prefixed: M = 0.6; β = 1.56, SE = 0.69, z = 2.27, *p* = 0.02). The same applies to externally prefixed forms, which were significantly easier to deal with in comparison to internally prefixed forms (Externally prefixed: M = 0.79; β = 1.27, SE = 0.62, z = 2.06, *p* = 0.04). Post hoc pairwise comparisons found no significant difference between externally prefixed and non-prefixed forms (OR = 1.34, SE = 1.06, z = 0.37, p = 0.72).

Verbs were easier to extract roots from compared to nouns (Verbs: M = 0.76; Nouns: M = 0.64; β = -0.77, SE = 0.14, z = 5.7, *p* < 0.001) and to gerunds (M = 0.7; β =

-0.64, SE = 0.22, z = -2.97, $p < 0.001$). Post hoc pairwise comparisons found no significant difference between nouns and gerunds (OR = 1.13, SE = 0.29, z = 0.49, p = 0.62).

Another generalized linear mixed model with random intercepts for participants and for items fit by maximum likelihood revealed that regular roots were significantly easier to extract compared to weak roots (Regular: M = 0.89; Weak: M = 0.72; β = -1.61, SE = 0.68, z = -2.35 , p = 0.02), and to quadriliteral roots (M = 0.73; β = -1.65, SE = 0.79, z = -2.08, p = 0.04) within the same patterns. Post hoc pairwise comparisons found no significant difference between weak and quadriliteral roots (OR = 0.96, SE = 0.86, z = -0.04, p = 0.97). Finally, a generalized linear mixed model with random intercepts for participants and for items fit by maximum likelihood found no significant difference between real words and pseudo-words (Real: M = 0.79; Pseudo: M = 0.74; β = 0.53, SE = 0.68, z = 0.78, p = 0.43).

3.3 Predictors of failure

We conducted a forward stepwise mixed-effects logistic regression with random intercepts for participants and for items to detect the factors that contributed to failure in root extraction. We chose this method to avoid multicollinearity by introducing variables one by one and checking if each new addition significantly improved the model's performance. One of our main goals in this analysis was to find the minimal interpretable model that explains failure. We hypothesized there would be various factors that make extracting the correct root more challenging based on the different variables that our experimental items contained, the challenges to root instruction and other common measures (e.g., frequency). The regression examined the contribution of each of these factors to failure in this task. The factors examined were:

- Lexical category (verb, noun and gerund)
- Root class (regular, weak and quadriliteral roots)
- Pseudo-roots versus real roots
- Occurrence in the verbal *CaCaC* pattern
- Infix vowels—the effect of different vowels as infixes
- Prefixation—prefixed versus non-prefixed forms
- Word frequency
- Lemma frequency
- Levenshtein distance from root—a calculation of the distance between two strings: the tested form and its root
- Levenshtein distance from past form—a calculation of the distance between two strings: the tested form and its past tense form

Pattern was excluded from this analysis as it conflates several factors that were modeled directly (e.g., infix vowel) and consequently does not offer independent predictive value.

The ordering of the factors presented in Table 4 reflects the extent to which each predictor improves model fit after stronger predictors have already been included.

Table 4. p and AIC values of the predictors of failure. Random intercepts for participants and for items were included

	Variable	*p*	ΔAIC
1	The vowel *i*	0.007	11
2	Prefixation	0.001	5.39
3	The vowel *u*	< 0.001	3.69
4	Lexical category	0.009	0.82
5	Root class	0.002	2.14
6	Levenshtein distance from past form	0.002	1.01

4. DISCUSSION

The general results showed a significant improvement in the children's ability to extract the consonantal skeleton of verbs, nouns, gerunds, and even pseudowords in Hebrew after participating in our program. Children were 258 times more likely to provide a correct response in the post-test compared to the pre-test, indicating a dramatic improvement following our intervention. The overall number of correct answers increased substantially, with a mean score of 0.73 in the post-test compared to 0.11 in the pre-test. Crucially, there was no difference in post-test accuracy between items that had appeared in both tests and items that appeared only in the post-test. In other words, prior exposure to an item did not provide an advantage, suggesting that the children acquired root extraction skills rather than relying on item-specific familiarity. This also suggests that our items were well matched in terms of difficulty across tests.

In the pre-test, second graders outperformed preschoolers and first graders, as expected, given their prior introduction to roots, and first graders also outperformed preschoolers. However, their scores were too low to indicate proficiency in root extraction (Second graders: M = 0.23; First graders: M = 0.12). Importantly, over half of the children (54.7%) scored 0 in the pre-test. This low performance can be attributed to several factors. First, the sample included children with differing levels of prior root knowledge. Second, root instruction in schools is often not sufficiently systematic for these skills to become fully internalized, even among children who have been introduced to them. By the post-test, not only had the overall performance improved substantially, but the effect of grade was no longer significant: preschoolers, first graders, and second graders performed similarly well. This convergence suggests that reading and writing skills are not essential for successful root instruction. Together, these findings address the first part of our first research question, demonstrating that the children improved their root extraction skills after participating in our program. To address the second part of the question, we conducted a linguistic analysis examining the children's ability to deal with words of different levels of difficulty.

It is important to note that approximately one third of the participants did not reach the final stage of the study. This attrition may have affected the findings, as it is likely that the children who completed all lessons and participated in the post-test were those who enjoyed the sessions and benefited most from the instruction. Such attrition is an inherent part of the experimental setting, which afforded limited control over children's engagement and procedures across families and relied primarily on parental feedback. Nevertheless, the dramatic improvement observed among the children who did complete the post-test indicates that, for at least two thirds of the participants, there was a clear and substantive internalization of the root-extraction method. Moreover, the significant differences between the pre- to post-test are consistent with results documented in two parallel experiments conducted in frontal, non-virtual instructional settings, both in Hebrew and in Arabic (Abu-Mukh et al., 2023; Gonen et al., 2025; Levinson et al., 2021, 2025). Even in the school-based setting, where all participants completed the intervention, a marked improvement was observed in all participating students' explicit morphological competence.

In the stepwise regression, gerunds and nouns showed lower predicted probabilities of success relative to verbs, making "lexical category" the fourth predictor of failure. Notably, this pattern emerged despite the fact that individual contrasts did not reach statistical significance in the reduced model presented in the linguistic analysis. A possible explanation for this can be that verbs are regularly manipulated in everyday speech, and while in verb inflection speakers perform various manipulations on the root, they do not necessarily do the same with nouns. Moreover, the fact that there are only seven verbal patterns as opposed to numerous nominal patterns may also facilitate extraction from verbs. Importantly, some verbal patterns were still more difficult for the children to deal with compared to other nominal patterns (see Appendix I).

Another challenge revealed in the linguistic analysis was the extraction of weak and quadriliteral roots, which were harder than regular ones. This factor emerged as the fifth most significant predictor of failure. All weak roots in the experiment were in the *CaCaC* pattern and transparent, i.e., recoverable from their past tense form. However, while trying to restore the missing root radical, 49.92% of the responses included the prefix instead of the verb's past form, which in this case was the correct response (e.g., providing *taʁad* for the form *t-eʁdu* "you(pl.) will go down" whose root and singular past form is *jaʁad*). The prefix was thus interpreted as a part of the root, leading to the creation of a novel "regular" root. This strategy was frequently employed in the children's responses even with regular roots, yielding novel quadriliteral roots with the prefix. Quadriliteral roots are generally less frequent, and 55.8% of erroneous responses involved their reduction to triliteral roots. Since the quadriliteral roots we used were reduplicated, (e.g., *l-χ-l-χ*), in 78.95% of these responses they were reduced to the common reduplicated form of triliteral roots (i.e., *l-χ-χ*). Nevertheless, the children successfully dealt with the effect of metathesis. When a root whose first radical is a strident appears in the pattern of *hitCaCeC*, it switches places with the *t* of the pattern, resulting in the pattern *hiStaCeC*. In these forms the *t* could be mistakenly interpreted as part of the root, yet the success rate of the metathical form we tested was as high as 95.3%.

The difficulty to deal with weak roots but not with metathical forms is related to the challenge posed by prefixation, which was also identified as the second most significant predictor of failure. Affixation in Hebrew serves both for inflection and derivation. In inflection, affixes can be either part of the pattern or a morphological realization of agreement features such as person, gender, number and tense, while in derivation they serve to create new words out of existing ones, often changing their lexical category (e.g., deriving an adjective from a noun) (Schwarzwald, 2018). It is often difficult to draw the line between an affix and a root radical, since they appear in a similar distribution (Deutsch et al., 2018; Schiff et al., 2020).

While examining the effect of prefixation on failure, we distinguished between prefixed and non-prefixed forms. Among the prefixed forms, we further distinguished between internally and externally prefixed forms. When the first root radical and the prefix share the same syllable, i.e., the first root radical is the coda of the prefix, the prefix is 'internal' (*t* in *ter.du*). When the prefix and the first root radical belong to two different syllables, the prefix is 'external' (*j* in *ji.ma.ʁaχ*). Internally prefixed words were significantly harder for children to extract roots from compared to externally and non-prefixed words, yet externally prefixed words were as challenging as non-prefixed words. The difficulty posed by internal prefixation likely stems from the shared distribution of the internal prefix and the first root radical. As reflected in the children's responses, internal prefixes were frequently interpreted as root radicals, particularly in weak roots and internally prefixed triliteral forms. In metathical forms the pattern's affix appears in a syllable of its own, effectively functioning as an external prefix and thus posing less of a challenge for extraction. Despite the challenges children faced with weak and quadriliteral roots, they still demonstrated high achievements and an ability to deal with these complex forms. We take this to indicate that our program contributed to their linguistic competence and an improved active morphological awareness.

An additional challenge that the children faced was extracting the roots of pseudowords, yet overall performance was high across both real words and pseudowords. Correspondingly, real-word status and form and lemma frequency were not found to contribute to failure, underscoring that word familiarity had little impact on success. Pseudowords are used in many neurolinguistic studies to examine phonological and morphological knowledge, specific language impairment (SLI), reading skills and dyslexia (Bar-On & Ravid, 2011; De Luca et al., 2002; Domahs et al., 2013; Maillart et al., 2004; Price et al., 1996). By using pseudowords we examined the children's morphological abilities while neutralizing intervening factors like word frequency and lexical-semantic knowledge, that often serve for analyzing real words. Pseudoword interpretation requires understanding the phonological constraints and morphological structure of Hebrew words, and the application of language rules to them (Argaman & Vaknin-Nusbaum, 2016; Ravid & Schiff, 2006).

Several studies on Hebrew inflectional and derivational morphology showed that children do better with real words as input rather than with pseudowords (Berman, 1993, 1999; Ravid & Schiff, 2006; Sokolov, 1988). These results were not fully replicated in our study as the children successfully dealt with pseudowords. Our lessons explicitly taught the children to extract the roots of pseudowords, a skill that is not part of schools' curricula. The successful analysis of pseudowords may again point to robust morphological awareness as well as reduced reliance on semantic relatedness during root extraction.

The presence of the vowels *i* and *u* as infixes in a pattern stood out as the first- and third-best predictors of failure, respectively. While the precise source of this impact of the vowel *i* is not entirely clear, we suggest it may stem from phonetic factors, specifically the height difference between *i* and the target vowel *a* of the *CaCaC* pattern. Future research will investigate whether this finding replicates in Arabic. In contrast, the source of the difficulty associated with the vowel *u* is more readily explained. The vowel *u* only appeared as an infix in the passive patterns *CuCaC* and *huCCaC*. Hebrew passive morphology is acquired very late and was found to be more challenging than other morphological constructions in studies (Ben-Zvi et al., 2025; Ravid & Vered, 2017).

Levenshtein distance from past form was identified as the sixth failure-predicting factors. Levenshtein distance is an algorithm that measures distances between words and has been used in linguistic research for comparing between dialects (Heeringa, 2004; Kessler, 1995) and evaluating genetic distance between languages (Serva & Petroni, 2008), inter alia. We calculated the Levenshtein distance of each form from its past tense form and its root by characters (i.e., the distance of a string like *jimaʁaχ* from the string of the past tense form *nimʁaχ* and from the string of the root *maʁaχ*). We found that the more distant the tested form was from its past tense form, the more difficult it was to extract the root correctly, and distance from the root itself was not significant. However, this difficulty may be a confound introduced by our teaching method. In the first pattern taught in the lessons, which is *CaCaC*, the root and the past tense form are identical. Hence, some of the children perceive the past tense form as the desired response, and in fact in many patterns the past tense form is closer to the stimulus than the root itself, since it contains prefixes too and/or has similar vowels (e.g., the form *jatχilu* whose 3SG past tense form is *hitχil* and root is *taχal*). This is in line with the fact that 28.7% of erroneous responses were such that provided the past tense form instead of the root (when applicable).

Surprisingly, "occurrence in *CaCaC*" was not found to be a significant predictor of failure. Among the seven verbal patterns of Hebrew, *CaCaC* is the most common one (Rubinstein, 2016). Nevertheless, there are roots that do not occur in this pattern at all. In our program, the children were taught to assign the root to the *CaCaC* pattern in their responses, whether it occurred in the pattern or not. Therefore, we assumed that roots that occurred in *CaCaC* would be easier to extract. These findings suggest that children do not rely on existing forms of the same root, in line with their limited reliance on semantic relatedness during root extraction.

To summarize this discussion, the results of our study clearly show that children significantly improved their root extraction skills following our online root instruction program. Furthermore, the children’s successful performance on complex forms points to a relatively advanced level of morphological competence. Finally, we found the most influential factors in children’s success in root extraction to be the infixes *i* and *u*, prefixation, lexical category, and root class. More precisely, the occurrence of the infixes *i* and *u* (and potentially passive morphology), non-verbal forms, and irregular roots extraction more challenging. These findings may be useful for both linguists and educational staff. Importantly, given the absence of a control group, the findings should be interpreted as evidence of significant pre–post improvement, rather than as proof of superiority over existing instructional approaches.

5. CONCLUSION

The results of the experiment show a substantial improvement in young children's root extraction abilities after participating in our program. Children extracted the roots of common verbs and also successfully handled more complex verbal forms, nouns and pseudowords. There was a drastic increase in the average success rate between the pre-test and the post-test, which soared from 0.11 to 0.73. There was no difference in accuracy between items that had appeared in both tests and items that appeared only in the post-test. Additionally, our program helped preliterate preschoolers to successfully learn root extraction and draw level with second graders, who had been introduced to roots in school. After having participated in our program, the significant difference that was found between the two groups in the pre-test was no longer present.

Since roots form an inherent part of Hebrew morphology and are relevant to many linguistic tasks such as novel word interpretation and literacy skills, learning to extract them is of high importance. Our program enhances this ability among young children, providing them with a deeper morphological awareness, and an improved linguistic competence. This is highly relevant to the educational system, as the program teaches skills that are necessary for language and high-school matriculation exams.

Our study further examined the predictors of failure in root extraction and found the presence of the vowels *i* and *u*, internal prefixation, non-verbality, and root irregularity to best predict failure in this task. We believe these findings to be important both on a linguistic and a pedagogical level. These findings provide a better understanding of the properties that make a word more challenging for root extraction and may be used to assist the educational staff in creating new, more effective curricula for root instruction while emphasizing the topics that were found to be more difficult for students.

Preliminary findings from a similar experiment where our program was adapted to teach root extraction to Arabic speaking children indicated similar improvement in root extraction skills (Abu-Mukh et al., 2023; Gonen et al., 2025). Future research will focus on providing a linguistic analysis for this study and compare our current findings to the ones found for Arabic. Even though our program is currently extracurricular, our method can be taught and implemented in schools in the form of videos, social robots or by teachers and can be used both in the Hebrew and Arabic curricula, for both first and second language instruction, in Israel and abroad.

ACKNOWLEDGEMENTS

Our research was supported by a grant from the Cukier-Goldstein-Goren Center for Mind, Cognition and Language (MILA) ([Goren Gordon and Einat Gonen 2019]). We would like to thank the children and their parents for participating in our program. We extend our gratitude to Genia Lukin and Shachar Ruppin for assistance with the statistical analyses.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the authors.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in OSF at https://osf.io/u57hv/overview?view_only=eb81ea104a274a84aaf9018fc3ca6061

REFERENCES

Abu-Mukh, A., Makhoul, B., Gonen, E., & Gordon, G. (2023). Social robots for teaching Arabic morphological awareness in elementary schools: A long-term in-the-wild study. In *ICERI2023 Proceedings*, 1001-1010. IATED. https://doi.org/10.21125/iceri.2023.0333

Akbulut, F. D. (2019). Role of morphological and metalinguistic awareness on reading among Turkish EFL learners. *International Journal of Contemporary Educational Research*, *6*(2), 261-277. https://doi.org/10.33200/ijcer.564134

Albert, A. (2019). The state of stop–fricative alternation in Modern Hebrew. *Brill's Journal of Afroasiatic Languages and Linguistics*, *11*(1), 135-161. https://doi.org/10.1163/18776930-01101010

Amirjalili, F., & Jabbari, A. A. (2018). The impact of morphological instruction on morphological awareness and reading comprehension of EFL learners. *Cogent Education*, *5*(1), 1523975. https://doi.org/10.1080/2331186X.2018.1523975

Argaman, O., & Vaknin-Nusbaum V. (2016). Language awareness of teacher trainees. *Journal of Language Teaching and Research, 7*(1), 9-18. https://doi.org/10.17507/jltr.0701.02

Bar-On, A., & Ravid, D. (2011). Morphological analysis in learning to read pseudowords in Hebrew. *Applied Psycholinguistics*, *32*(3), 553-581. https://doi.org/10.1017/S014271641100021X

Bentin, S., & Frost, R. (2013). Morphological factors in visual word identification in 1 2 Hebrew. *Morphological Aspects of Language Processing*, 271–292.

Ben-Zvi, G., Landau, H., & Ravid, D. (2025). A developmental study of Hebrew text reconstruction: The role of passive voice in different genres. *First Language, 45(4),* 496–532.

https://doi.org/10.1177/01427237251339787

Berent, I., & Shimron, J. (1997). The representation of Hebrew words: Evidence from the contour principle. *Cognition, 64*, 39–72. https://doi.org/10.1016/S0010-0277(97)00016-4

Berman, R. A., & Bolozky, S. (1978). *Modern Hebrew structure* (p. 164). University Pub. Projects. https://doi.org/10.2307/413968

Berman, R. A. (1993). Developmental perspectives on transitivity: A confluence of cues. In Y. Levy (Ed.), *Other Children, Other Languages: Issues in the Theory of Acquisition*, 189–241.

Berman, R. A. (1999). Children's innovative verbs versus nouns: Structured elicitations and spontaneous coinages. In L. Menn & N. Bernstein-Ratner (Eds.), *Methods in Studying Language* Production (pp. 69–93).

Berman, R. A. (2003). Children's lexical innovations: Developmental perspectives on Hebrew verb structure. *Language Processing and Acquisition in Languages of Semitic, Root-based Morphology*, *28*, 243-292. https://doi.org/10.1075/lald.28.13ber

Bolozky, S. (1999). Measuring productivity in word formation: the case of Israeli Hebrew. *Studies in Semitic Languages and Linguistics 27.* Brill. https://doi.org/10.1163/9789004348431

Bolozky, S. (2003). The 'roots' of denominative Hebrew verbs. *Language Processing and Acquisition in Languages of Semitic, Root-based Morphology*, 131-146. https://doi.org/10.1075/lald.28.07bol

Boudelaa, S. (2014). Is the Arabic mental lexicon morpheme-based or stem-based? Implications for spoken and written word recognition. *Handbook of Arabic Literacy*, 31-54. https://doi.org/10.1007/978-94-017-8545-7_2

Bowers, P. N., Kirby, J. R., & Deacon, S. H. (2010). The effects of morphological instruction on literacy skills: A systematic review of the literature. *Review of Educational Research*, *80*(2), 144-179. https://doi.org/10.3102/0034654309359353

Carlisle, J. F., & Feldman, L. B. (1995). Morphological awareness and early reading achievement. *Morphological Aspects of Language Processing*, 189-209.

Carlisle, J. F. (2010). Effects of instruction in morphological awareness on literacy achievement: An integrative review. *Reading Research Quarterly*, *45*(4), 464-487. https://doi.org/10.1598/RRQ.45.4.5

Central Bureau of Statistics (2023). Compatibility to the subject among teaching staff in Hebrew, mathematics and English, 2012-2022. https://www.cbs.gov.il/he/mediarelease/DocLib/2023/049/06_23_049b.pdf

Clark, E. V., & Berman, R. A. (1984). Structure and use in the acquisition of word formation. *Language*, 542-590. https://doi.org/10.2307/413991

Darling-Hammond, L., DiNapoli Jr, M., & Kini, T. (2023). The federal role in ending teacher shortages. *Learning Policy Institute*. https://doi.org/10.54300/649.892

De Luca, M., Borrelli, M., Judica, A., Spinelli, D., & Zoccolotti, P. (2002). Reading words and pseudowords: An eye movement study of developmental dyslexia. *Brain and Language*, *80*(3), 617-626. https://doi.org/10.1006/brln.2001.2637

Derwing, B. L., Smith, M. L., & Wiebe, G. E. (1995). On the role of spelling in morpheme recognition: Experimental studies with children and adults. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 3–27).

Deutsch, A., Velan, H., & Michaly, T. (2018). Decomposition in a non-concatenated morphological structure involves more than just the roots: Evidence from fast priming. *Quarterly Journal of Experimental Psychology*, *71*(1), 85-92. https://doi.org/10.1080/17470218.2016.1250788

Deutsch, A., & Kuperman, V. (2019). Formal and semantic effects of morphological families on word recognition in Hebrew. *Language, Cognition and Neuroscience*, *34*(1), 87-100. https://doi.org/10.1080/23273798.2018.1513541

Domahs, U., Lohmann, K., Moritz, N., & Kauschke, C. (2013). The acquisition of prosodic constraints on derivational morphology in typically developing children and children with SLI. *Clinical Linguistics & Phonetics*, *27*(8), 555-573. https://doi.org/10.3109/02699206.2013.796405

Donitsa-Schmidt, S. (2025) Teacher shortage in Israel: A policy analysis of recruitment, preparation, and retention. *Journal of Education for Teaching*, *51*(5), 847-865. https://doi.org/10.1080/02607476.2025.2563153

Faust, N. (2019). New reasons to root for the Semitic root from Mehri and Neo-Aramaic. *The Linguistic Review, 36*(3), 575-599. https://doi.org/10.1515/tlr-2019-2030

Flores, M. A., and C. J. Craig. 2023. Reimagining teacher education in light of the teacher shortage and the aftershock of COVID-19: Adjusting to a rapidly shifting world. *European Journal of Teacher Education*, *46*(5), 772–788. https://doi.org/10.1080/02619768.2023.2294697

Frost, R., Forster, K. I., & Deutsch, A. (1997). What can we learn from the morphology of Hebrew? A masked-priming investigation of morphological representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(4), 829. https://doi.org/10.1037//0278-7393.23.4.829

Frost, R., Deutsch, A., & Forster, K. I. (2000). Decomposing morphologically complex words in nonlinear morphology. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 751–765. https://doi.org/10.1037//0278-7393.26.3.751

Gonen, E., Abu-Mukh, A., & Gordon, G. (2025). Socially sustainable HRI for underrepresented communities: Case study of language learning of Arab children in Israel. In *2025 20th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 361-370. IEEE. https://doi.org/10.1109/hri61500.2025.10974026

Heeringa, W. J. (2004). Measuring dialect pronunciation differences using Levenshtein distance. https://pure.rug.nl/ws/portalfiles/portal/9800656/thesis.pdf

Jakubíček, M., Kilgarriff, A., Kovář, V., Rychlý, P., & Suchomel, V. (2013). The TenTen corpus family. In *7th International Corpus Linguistics Conference CL* (pp.125-127). https://doi.org/10.1007/s40607-014-0009-9

Kessler, B. (1995). Computational dialectology in Irish Gaelic. In *Seventh conference of the European Chapter of the Association for Computational Linguistics*. https://doi.org/10.3115/976973.976983

Kirk, C., & Gillon, G. T. (2009). Integrated morphological awareness intervention as a tool for improving literacy. *Language, Speech & Hearing Services in Schools, 40(3)*, 341-351. https://doi.org/10.1044/0161-1461(2008/08-0009)

Levie, R., Dattner, E., Zwilling, R., Rosenstein, H., Stanzas, S. E., & Ravid, D. (2019). Complexity and density of Hebrew verbs in preschool peer talk: The effect of socio-economic background. *The Mental Lexicon*, *14*(2), 237-273. https://doi.org/10.1075/ml.00006.lev

Levin, I., Ravid, D., & Rapaport, S. (1999). Developing morphological awareness and learning to write: A two-way street. In *Learning to read: An integrated view from research and practice* (pp. 77-104). Dordrecht. https://doi.org/10.1007/978-94-011-4826-9_5

Levin, I., Ravid, D., & Rapaport, S. (2001). Morphology and spelling among Hebrew-speaking children: From kindergarten to first grade. *Journal of Child Language*, *28*(3), 741-772. https://doi.org/10.1017/S0305000901004834

Levin, I., Saiegh-Haddad, E., Hende, N., & Ziv, M. (2008). Early literacy in Arabic: An intervention study among Israeli Palestinian kindergartners. *Applied Psycholinguistics*, *29*(3), 413-436. https://doi.org/10.1017/S0142716408080193

Levinson, L., Gvirsman, O., Gorodesky, I. M., Perez, A., Gonen, E., & Gordon, G. (2021). Learning in summer camp with social robots: A morphological study. *International Journal of Social Robotics*, *13*(5), 999-1012. https://doi.org/10.1007/s12369-020-00689-y

Levinson, L., Melamed Gorodesky, I., Michaelovski, V., Gordon, G., & Gonen, E. (2025). Hebrew roots and patterns: Developing early childhood morphological knowledge through an oral teaching methodology & social robots. *Innovation in Language Learning and Teaching*, 1-18. https://doi.org/10.1080/17501229.2024.2401475

Lipkin, S. (1985). *Modern Hebrew phonology for junior high school*. CET.

Maillart, C., Schelstraete, M. A., & Hupet, M. (2004). Phonological representations in children with SLI. *Journal of Speech, Language, and Hearing Research*, *47*(1), 187–198. https://doi.org/10.1044/1092-4388(2004/016)

Marslen-Wilson, W. (2016). Abstractness and combination: The morphemic lexicon. In *Language processing* (pp. 101-119). Psychology Press.

McCarthy, J. J. (1981). A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry*, *12*(3), 373-418. https://doi.org/10.2307/4178229

Mimouni, Z., Kehayia, E., & Jarema, G. (1998). The mental representation of singular and plural nouns in Algerian Arabic as revealed through auditory priming in agrammatic aphasic patients. *Brain and Language*, *61*(1), 63-87. https://doi.org/10.1006/brln.1997.1879

Price, C. J., Wise, R. J. S., & Frackowiak, R. S. J. (1996). Demonstrating the implicit processing of visually presented words and pseudowords. *Cerebral Cortex*, *6*(1), 62-70. https://doi.org/10.1093/cercor/6.1.62

Ravid, D., Avivi, G., & Levy, R. (1999). Derivational morphology in SLI children: Structure and semantics of Hebrew nouns. In *New directions in language development and disorders* (pp. 39-49). Springer US. https://doi.org/10.1007/978-1-4615-4157-8_5

Ravid, D. (2003). A developmental perspective on root perception in Hebrew and Palestinian Arabic. *Language Acquisition and Language Disorders*, *28*, 293-320. https://doi.org/10.1075/lald.28.14rav

Ravid, D., Ashkenazi, O., Levie, R., Zadok, G. B., Grunwald, T., Bratslavsky, R., & Gillis, S. (2016). Foundations of the early root category. *Acquisition and Development of Hebrew: From Infancy to Adolescence*, *19*, 95-134. https://doi.org/10.1075/tilar.19.04rav

Ravid, D., & Bar-On, A. (2001). The Semitic root in language acquisition. *GALA (Generative Approaches to Language Acquisition), 2001 Proceedings*.

Ravid, D., & Malenky, A. (2001). Awareness of linear and nonlinear morphology in Hebrew: A developmental study. *First Language*, *21*(61), 25-56. https://doi.org/10.1177/014272370102106102

Ravid, D., & Schiff, R. (2006). Roots and patterns in Hebrew language development: Evidence from written morphological analogies. *Reading and Writing*, *19*(8), 789-818. https://doi.org/10.1007/s11145-006-9004-3

Ravid, D., & Vered, L. (2017). Hebrew verbal passives in later language development: The interface of register and verb morphology. *Journal of Child Language*, *44*(6), 1309-1336. https://doi.org/10.1017/S0305000916000544

Rimon, N., Gorodesky, I. M., Gordon, G., & Gonen, E. (2022). Child's play: Teaching young children Hebrew morphology using E-learning technologies. *International Association for Development of the Information Society*. https://doi.org/10.33965/es_ml2022_202202c036

Rosner, R. (2003). *Influence of the development of Hebrew language research - morphology - on Hebrew grammar text books for high school in the period of 1948-2000* [Doctoral dissertation, Hebrew University of Jerusalem].

Rubinstein, D. (2016). Remarks on the vocabulary of the corpus. In E Gonen (ed.), *Studies in Spoken Hebrew*, *27*, 129-149.

Schiff, R., Raveh, M., & Kahta, S. (2008). The developing mental lexicon: Evidence from morphological priming of irregular Hebrew forms. *Reading and Writing*, *21*(7), 719-743. https://doi.org/10.1007/s11145-007-9088-4

Schiff, R., & Saiegh-Haddad, E. (2018). Development and relationships between phonological awareness, morphological awareness and word reading in spoken and standard Arabic. *Frontiers in Psychology*, *9*, 356. https://doi.org/10.3389/fpsyg.2018.00356

Schiff, R., Rosenstock, S., & Ravid, D. (2020). Morpho-orthographic complexity in affix spelling in Hebrew: A novel psycholinguistic outlook across the school years. *Frontiers in Psychology*, *11*, 868. https://doi.org/10.3389/fpsyg.2020.00868

Schwarzwald, O. R. (2018). Stress assignment in words with-i suffix in Hebrew. *SKASE Journal of Theoretical Linguistics*, *15*(3), 110-127.

Serva, M., & Petroni, F. (2008). Indo-European languages tree by Levenshtein distance. *Europhysics Letters*, *81*(6), 68005. https://doi.org/10.1209/0295-5075/81/68005

Sokolov, J. L. (1988). Cue validity in Hebrew sentence comprehension. *Journal of Child Language, 15*(1), 129-155. https://doi.org/10.1017/S0305000900012095

Taha, H., & Saiegh-Haddad, E. (2017). Morphology and spelling in Arabic: Development and interface. *Journal of Psycholinguistic Research*, *46*(1), 27-38. https://doi.org/10.1007/s10936-016-9425-3

Temkin Martínez, M. (2010). *Sources of non-conformity in phonology: Variation and exceptionality in Modern Hebrew spirantization* [Doctoral dissertation, Boise State University]. https://scholarworks.boisestate.edu/linguistics/4/

Vaknin-Nusbaum, V., Sarid, M., Raveh, M., & Nevo, E. (2016). The contribution of morphological awareness to reading comprehension in early stages of reading. *Reading and Writing*, *29*, 1915-1934. https://doi.org/10.1007/s11145-016-9658-4

Vaknin-Nusbaum, V., & Raveh, M. (2019). Cultivating morphological awareness improves reading skills in fifth-grade Hebrew readers. *The Journal of Educational Research*, *112*(3), 357-366. https://doi.org/10.1080/00220671.2018.1528541

Vaknin-Nusbaum, V. (2021). Morphological awareness: A tool to promote reading fluency and accuracy in Hebrew in students with reading disabilities. *Learning Disabilities Research & Practice*, *36*(1), 26-39. https://doi.org/10.1111/ldrp.12236

Yucel-Koc, M. (2015). *The role of morphological awareness in academic vocabulary and reading comprehension skills of adult ESL learners*. Seattle Pacific University.

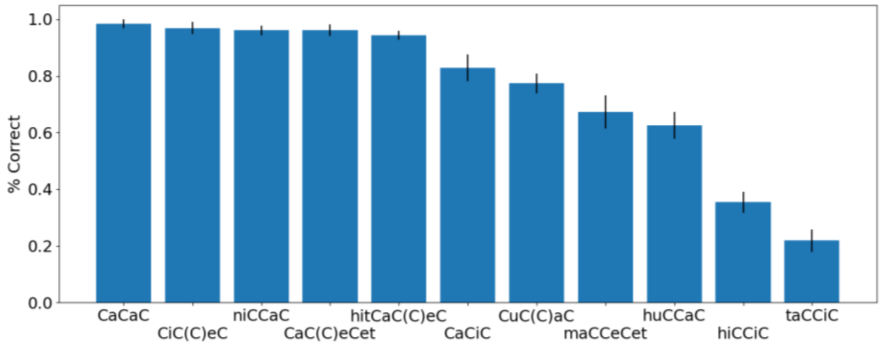
APPENDIX 1: PATTERNS

In order to further assess the challenges in root extraction, we made a comparison between the different patterns examined. To neutralize confounding factors, we removed all pseudowords and words with weak or quadriliteral roots (all weak roots used belonged to the *CaCaC* pattern and drastically lowered its score. Quadriliteral roots naturally appear in only three constant patterns, of which only two were used). A one-way repeated measures ANOVA found a significant main effect for pattern ($F(1,63) = 3237.54$, $p < 0.001$). Post-hoc Bonferroni-corrected analyses showed that *hiCCiC* and *taCCiC* were significantly harder than all other patterns, having *taCCiC* significantly harder than *hiCCiC*.

Table A.1. Mean score and standard deviation of each pattern

Pattern	Pattern Type	Mean (SD)
CaCaC	Verbal	0.98 (0.02)
CiC(C)eC	Verbal	0.97 (0.02)
niCCaC	Verbal	0.96 (0.02)
CaC(C)eCet	Nominal	0.96 (0.02)
hitCaC(C)eC	Verbal	0.94 (0.02)
CaCiC	Nominal	0.83 (0.05)
CuC(C)aC	Verbal	0.77 (0.04)
maCCeCet	Nominal	0.67 (0.06)
huCCaC	Verbal	0.63 (0.05)
hiCCiC	Verbal	0.35 (0.04)
taCCiC	Nominal	0.22 (0.04)

Figure A.1. Mean score and standard deviation of each pattern (Error bars represent ±1 standard error)



The results suggest a two-group division for the remaining patterns:

- Group I: CaCaC, CiC(C)eC, niCCaC, CaC(C)eCet, hitCaC(C)eC
- Group II: CaCiC, CuC(C)aC, maCCeCet, huCCaC

All patterns of Group I were significantly easier than all patterns of Group II. There was no significant difference between the patterns within Group I. However, in Group II, *CaCiC* was found to be significantly easier than *maCCeCet* and *huCCaC*. Other comparisons did not result in a significant difference.

Relying on the predictors of failure that were found, we suggest that the patterns in Group I were significantly easier to extract roots from since none of their forms had an internal prefix. It is important to note that the forms we tested in *niCCaC* were all future inflected, which is where the prefix is external. The pattern's characteristic internal prefix *ni-* appears only in the past and present tense forms, yet such forms were not inspected. Given findings from our current and previous studies, such forms were supposed to exhibit a much lower success rate, and future research demands taking them into account for a more coherent picture.

In Group II, the first two patterns had a high vowel as an infix, and the last two had an internal prefix. *CaCiC* was significantly easier than *maCCeCet* and *huCCaC*, probably because it is neither suffixed nor (internally) prefixed.

APPENDIX 2: PARENTS' FEEDBACK

Since our program was extracurricular, parents were very engaged in the program as they had to play the videos for the children and then submit feedback forms for them. Consequently, parents were familiar with the program's curriculum and reported that they offered the children to play "the root extraction game" in their free time and were even offered to do so by the children. In the final feedback form, parents rated their agreement with several statements on a scale from 1 to 5, while 1 indicated “strongly disagree” and 5 indicated “strongly agree”. The responses presented in the table below indicate that parents were satisfied with our program and believed it had great impact on their children’s linguistic competence.

Table A.2. Parents' mean ratings on the final feedback questionnaire (1 = strongly disagree, 5 = strongly agree)

Statement	Mean rating
My child learned a lot in the program	4.14
My child enjoyed the lessons	4.1
I would like my children to have more digital educational contents at school/preschool	4.01
I would like my children to have more digital educational contents at home	4.1
The ability to extract roots is important for children's linguistic competence	4.28
We would like to participate in similar programs in the future	4.5

APPENDIX 3: A PARTIAL TRANSCRIPTION OF A LESSON

Hi! My name is Fuzzy. Yes, Fuzzy! You probably want to know what I am. I'm a yellow piece of fur! Hahaha. I really wanted to come visit your school. I hope you could meet my friends later. My friend Tutty and I started the "Root Detective" club. We would be very happy if you joined us. Would you like to join us? In order to join our club, you need to know roots of words. No! Not roots of trees, roots of words! Hahaha.

Roots are the base of word families. For example:
liʁkod ("to dance") – *ʁakad*
lismoaχ ("to be happy") – *samaχ*
liv'ot ("to kick") – *ba'at*
linʃom ("to breathe") – *naʃam*
ligdol ("to grow") – *gadal*
lispoʁ ("to count") – *safaʁ*

Yes, it has a rhythm! Maybe we should start a band! Hahaha. Now it's your turn!

What is the root of *liχtov* ("to write")?

-*katav*. Excellent!

What is the root of *laχʃov* ("to think")?

-*xaʃav*. Correct!

What is the root of *linhog* ("to drive")?

-*nahag*. Great!

What is the root of *lisgoʁ* ("to close")?

-*sagaʁ*. Wonderful!

What is the root of *liχ'os* ("to be angry")?

-*ka'as*. Amazing!

What is the root of *la'amod* ("to stand up")?

-*amad*. Well done!

How do you know all the answers?

…

Oh, I almost forgot! We really like to play, and we especially like to stand up and jump, but I'm just a cute yellow creature with furry legs, Hahaha.

Hey! Maybe you like to play too?

Let's stand up quick quick quick.

Now, let’s jump. Everybody jump! Jump jump jump!
What is the root of *likfot͡s* (“to jump”)?
-*kafat͡s*. Correct!
And now let’s step on our left foot, and now on our right foot.
What is the root of *nidʁoχ* (“let’s step”)?
-*daʁaχ*. Let’s march!
What is the root of *nit͡s’ad* (“let’s march”)?
- *t͡sa’ad*. Excellent!
Now let’s go back to our seats and sit down.
What is the root of *χozʁim* (“we are going back”)?
-*χazaʁ*. That is correct! Great job!