Morphology (2020) 30:1–60
https://doi.org/10.1007/s11525-020-09348-4

The route to the derivational verb family in Hebrew: A psycholinguistic study of acquisition and development

Ronit Levie[1] · Orit Ashkenazi[1] · Shirly Eitan Stanzas[1] · Rachel (Chelli) Zwilling[1] · Efrat Raz[1] · Liat Hershkovitz[1] · Dorit Ravid[1]

Received: 19 May 2019 / Accepted: 16 January 2020 / Published online: 6 February 2020

Abstract A main challenge for language users is forging reliable relationships between words with shared components so that morphology as a system emerges from usage. For native Hebrew acquisition this means that learners acquire verbs as lexical entities, which form into a system based on Semitic roots and *binyan* conjugations. The Hebrew verb system is consequently organized by derivational families, where verbs in different *binyan* conjugations share the same root. This is illustrated by the *k-t-b*-based family containing *katav* 'write', *nixtav* 'be written', *hixtiv* 'dictate', *huxtav* 'be dictated', *kitev* 'carbon copy [cc]', *kutav* 'be cc'ed', and *hitkatev* 'correspond'. The current study offers a systematic account of how Hebrew verb families and their components—verb lemmas, roots and *binyan* patterns—emerge and develop in structural and semantic terms, covering the long route from infancy to adulthood. The study is grounded in a large database (485,908 word tokens) compiled of the spoken and written productions of Hebrew-speaking toddlers, children, adolescents and adults.

The study is presented in two parts. Part I describes the general characteristics of the study database with regards to the distributions of verbs, roots and *binyan* verb conjugations, focusing on developmental changes as indicators of the growth and consolidation of the verb lexicon. Part II presents the development of root-based verb derivational families in terms of family frequency, family size, family composition and the semantic coherence of families. Based on the empirical evidence in this paper, our main claim is that roots, *binyan* conjugations and derivational verb families are all *emergent properties* of the verb system as it develops in variegated communicative contexts.

✉ R. Levie
levieron@gmail.com

D. Ravid
doritr@tauex.tau.ac.il

1 Tel Aviv University, Tel Aviv, Israel

Keywords Hebrew · Derivation · Verbs · Root · *binyan* · Development

Part I

1 Introduction

Derivational morphology organizes the mental lexicons of many languages (Bybee 1988; Haspelmath and Sims 2013; Marslen-Wilson 2007; Paterson et al. 2011). In morphology-rich languages such as Hebrew, where many grammatical and lexical notions are encoded in word-internal structures (Deutsch and Kuperman 2019; Kastner 2019; Ravid 2006, 2012), gaining command of derivational morphological devices is paramount for the acquisition and processing of the lexicon. The current study takes a developmental psycholinguistic and typological perspective on the morphology of Hebrew derivational verb families, the major habitat of what is known as Semitic *root-and-pattern*, or *non-linear* morphology (Boudelaa and Marslen-Wilson 2005; McCarthy 1981).

While other lexical classes also partially rely on root-and-pattern morphology (Berman 1987; Deutsch and Malinovitch 2016; Ravid 1990, 2006), the Hebrew derivational verb system is completely non-linear.[1] Moreover, temporal stem structure in verbs is also the only non-linear inflectional system in the language (Ravid and Malenky 2001; Schwarzwald 2002). The Hebrew verb-pattern system is at the same time the first derivational class learned in early childhood and the prototype exemplar of a non-linear system (Berman 1985; Ben Zvi and Levie 2016).

The development of Semitic verb morphology in native speakers has long challenged the psycholinguistic literature (Berman 1987, 2012; Ravid 2003, 2012). Developmental accounts need to explain how Hebrew verbs, lexical entities, are learned in a language where root and *binyan* verb-patterns[2]—sub-lexical, discontinuous, unpronounceable morphemes—are critical components of verb structure and meaning. Such accounts also need to explain the emergence and consolidation of root-based derivational verb families, with a single root shared by verbs with different *binyan* conjugations, such as *katav* 'write', *nixtav* 'be written', *hixtiv* 'dictate', *huxtav* 'be dictated', *kitev* 'carbon copy [cc]', *kutav* 'be cc'ed', and *hitkatev* 'correspond'—all

[1]The well-motivated consensus is that *binyan* prefixes such as *Nif'al n-* or *Hitpa'el hit-* are part of the *binyan* pattern (Bolozky 2007; Schwarzwald 2002). This consensus is supported by children's usage patterns that do not reflect any separate status for the prefix (Berman 1981; Ravid 1995).

[2]Two separate terms are used here for the notion of *binyan*. The term 'pattern' or '*binyan* pattern' is used to refer here to a morpho-phonological template uniquely associated with a temporal category in a specific *binyan*. For example, *CaCaC* and *CoCeC* are patterns respectively associated with the past and present tense categories of the *Qal* conjugation, as further elaborated in the introduction (and also see Ravid 2019). In contrast, the term 'conjugation' is more abstract. It is reserved to the unique set of temporal patterns making up a single *binyan* (e.g., both *Qal* and *Hif'il* constitute *binyan* conjugations, each consisting of a unique set of temporal patterns). Accordingly, a verb lemma is determined as a specific combination of a root and a *binyan* conjugation, e.g.,'learn' as the unique combination of *l-m-d* and *Qal*. This lemma can occur in five different templates representing the temporal *Qal* patterns.

based on root *k-t-b* 'write'. Derivational families organize the verb lexicon morphophonologically as well as in terms of transitivity relations (and other semantico-syntactic relations, as elaborated in Ravid 2019). They are thus critical for both morphological and syntactic acquisition. This organization has served as the topic of several studies (Armon-Lotem and Berman 2003; Ashkenazi et al. 2016, 2019; Berman 1985, 1993a, 1993b, 2000, 2003; Levie et al. 2019; Lustigman 2013; Ninio 1999; Ravid et al. 2016). But to date, no study has offered a systematic account of how Hebrew verb families and their components—verb lemmas, roots and *binyan* patterns—emerge and develop in structural and semantic terms, covering the long route from infancy to adulthood. This is the goal of the current study, grounded in a large set of new corpora compiled of the spoken and written productions of Hebrew-speaking toddlers, children, adolescents and adults (henceforth termed 'the compiled database').

The general conceptual framework of the current study is the Usage Based approach to linguistics and psycholinguistics, according to which speakers construct grammatical systematicity from experience with individual usage events in a process that is graded, probabilistic, interactive, context-sensitive and domain-general (Goldberg 2006; Tomasello 2003). Recent usage-based accounts of morphological learning, use and change (specifically expressed in the *word-and-paradigm* approach) have turned towards the word as the fundamental unit in morphology (Ackerman et al. 2009; Blevins 2016; Bonami and Stump 2016; Traugott and Trousdale 2013). In this view, the main challenge for language users is to forge reliable relationships between words with shared components so that morphology as a system emerges from usage (Abbot-Smith and Tomasello 2006; Ackerman and Malouf 2013; McCauley and Christiansen 2019).

Due to the large scope of the current paper, it is presented in two parts—Part I and Part II. Part I below describes the general characteristics of the study database with regards to the distributions of verbs, roots and *binyan* verb conjugations, focusing on developmental changes as indicators of the growth and consolidation of the verb lexicon. It consists of Sects. 1–5, including the general introduction to Hebrew verb morphology in terms of structure and semantics (Sect. 1), the aims and hypotheses of the study (Sect. 2), the methods (Sect. 3), results (Sect. 4) and discussion (Sect. 5), ending with an interim conclusion.

Part II following Part I presents the development of root-based verb derivational families in terms of family frequency, family size, family composition and the semantic coherence of families. It consists of Sect. 6 on the results and discussion regarding these four facets of root-based derivational families, and the final concluding discussion of the paper (Sect. 7).

1.1 Root-related derivational verb families

Non-linear morphology involves a densely related verb lexicon both semantically and structurally. Consider example (1a–c) below, demonstrating the non-linear root and pattern make up of Hebrew verbs: Unlike English, where semantically related verbs largely take different lexical forms,[3] corresponding Hebrew verbs are structurally related as well, as shown in (1).

[3]This is not to say that English does not relate verbs via morphological means (vowel change as in *sit/set/seat*; negative prefixation as in *do/undo*, not to mention compound constructions such as *brain-*

(1)	1a	*katav*	*hixtiv*	*kitev*	*hitkatev*	*k-t-b*
		write	dictate	(to) cc	correspond	
	1b	*gadal*	*higdil*	*gidel*	*hitgadel*	*g-d-l*
		grow	enlarge	raise	self-aggrandize	
	1c	*kadam*	*hikdim*	*kidem*	*hitkadem*	*q-d-m*[4]
		precede	come early	promote	move forward	

As example (1) shows, Hebrew verbs are related in two ways, presented in this example along the horizontal (root-related) and vertical (*binyan*-related, see Sect. 1.2) axes. Horizontally, a tri-radical root is shared as the consonantal skeleton of four verbs on the same line. For example, root *g-d-l* is shared by all verbs in (1b). These sets of root-related verb lemmas are termed a *root-based derivational family* (Blevins 2014). Most Hebrew morphologists regard the shared root skeleton as carrying a basic, shared meaning typical of the derivational family, such as ‘write’ in the case of *k-t-b*, ‘increase’ for *g-d-l*, or ‘come early’ for *q-d-m* (Berman 1987; Bolozky 1999; Kastner 2019; Laks 2013; Schwarzwald 2000). Developmental studies point to an early ability of Hebrew-speaking children to extract roots from familiar words and use them in novel forms (Berman 1985, 2000, 2012; Berman and Sagi 1981; Clark 2003; Ravid 2003), such as juvenile *nigdal* ‘grow’ (cf. conventional *gadal*). Current evidence points to the Semitic root as the most accessible Hebrew morpheme in spoken and written language development (Ben Zvi and Levie 2016; Gillis and Ravid 2006; Ravid 2001, 2019; Ravid and Bar On 2005; Seroussi 2011), even in contexts of language disability or environmental deprivation (Levie et al. 2017, 2019; Ravid et al. 2003; Ravid and Schiff 2006a; Schiff and Ravid 2007). Reading and spelling research also demonstrates that Hebrew words are linked through their roots (Bar-On and Ravid 2011; Deutsch and Meir 2011; Frost 2012; Frost et al. 2000; Ravid 2012; Ravid and Schiff 2006b; Schwarzwald 1981; Velan et al. 2005). This is the basis for the prevalent view of the Semitic root as the lexical core of Hebrew words, and in particular of verbs.

A different, phonology-oriented approach to Hebrew verb structure, termed “stem-based” or “word-based”, assumes that there is no morphemic consonantal root at the base of the derivation (Bat-El 1994, 2003, 2017; Ussishkin 2005). This approach denies the existence of the root as a morpheme, treating it as epiphenomenal to the morphemic template of the pattern (*binyan*); the latter is indeed granted the status of a morpheme. Specifically, this approach derives all verbs from a base form of *CaCaC*, i.e., the citation form of the *Qal* pattern. In Sect. 7.1 below we discuss the stem/word-based approach in the context of the morphological, syntactic and semantic acquisitional evidence presented in the current study.

wash), however we argue that these are not comparable to the single, mandatory device of root and *binyan* morphology in Hebrew.

[4]Roots are represented as morphological entities, that is, taking into account their morpho-phonological behavior, as detailed in Ravid (2012). For example, the *k* in root *k-t-b* alternates with spirant *x*, while the *k* in *q-d-m* does not (Temkin Martínez 2010). For words (in contrast to roots) we use a broad phonemic transcription.

1.2 *Binyan* conjugations in derivational families

Example (1) shows that Hebrew verbs share another, vertically oriented, relationship, in addition to the horizontal root-based one (Levie et al. 2019). Note that across the three derivational families in (1a–c), verbs based on different roots share similar stems constructed by vocalic structures that complement the consonantal root. For example, *kitev* 'cc', *gidel* 'raise', and *kidem* 'promote' all share the stem form *CiCeC*, with C's standing for root radicals. In the same way, *hixtiv* 'dictate', *higdil* 'enlarge', and *hikdim* 'come early' share the form *hiCCiC*. These root-complementing morphemes, termed *verb patterns*, are the vocalic templates within which root consonants are couched. Patterns provide the stem vowels and prosodic template, including the specific sites where root radicals intersperse with vowels, as well as prefixes in some cases. Thus, they in fact determine the basic morpho-phonology of the verb stem.[5]

In the traditional sense, the notion of verb pattern is taken to refer to seven conjugations termed *binyanim* (literally, 'buildings')—named *Qal* (*Pa'al*), *Nif'al, Hif'il, Huf'al, Pi'el, Pu'al*, and *Hitpa'el*.[6] In verb derivational families, verbs sharing the same root are based on different *binyan* conjugations, as shown in (1a–c) above. The size of a derivational verb family in fact indicates how many *binyan* conjugations are assigned to the same root. Derivational families[7] range from singleton verbs, with no root-related family members, to larger families of two up to seven members. Example (1a–c) above illustrated three derivational verb families, each composed of four members.[8]

(2)

Binyan / Family size	*Qal*	*Nif'al*	*Hif'il*	*Pi'el*	*Hitpa'el*
Singleton				*tipes* climb	
Two members		*nirdam* fall asleep	*hirdim* put to sleep		
Three members	*asaf* collect	*ne'esaf* be collected, gather, Int			*hit'asef* gather, Int
Five members	*yada* know	*noda* become known	*hodi'a* announce	*yidé'a* inform	*hitvada* become acquainted

[5]By "the basic morpho-phonology of the verb stem" we mainly refer to the stem vowels and its prosodic structure, but also to the morpho-phonological contexts of *bkp* stop/spirant alternation (Kastner 2019; Ravid 1995).

[6]The traditional names of the *binyan* conjugations actually refer to their past tense patterns, with root *p-'-l* 'act, do' filling in the site of root radicals.

[7]As one of our readers pointed out, a family needs more than one member. We use this term to refer to the nature of root-based verbs, which may or may not be related via the same root.

[8]The semi-automatic passive conjugations *Huf'al* and *Pu'al*, corresponding to transitive *Hif'il* and *Pi'el* respectively, are not included in this example.

Table 1 The seven *binyan* conjugations as sets of temporal patterns

Binyan	Past Tense	Present Tense	Future Tense	Imperative	Infinitive
Qal	*CaCaC*	*CoCeC*	*yiCCoC*	*CCoC*	*liCCoC*
Nif'al	*niCCaC*	*niCCaC*	*yiCaCeC*	*hiCaCeC*	*lehiCaCeC*
Hif'il	*hiCCiC*	*maCCiC*	*yaCCiC*	*haCCeC*	*lehaCCiC*
Huf'al	*huCCaC*	*muCCaC*	*yuCCaC*	–	–
Pi'el	*CiCeC*	*meCaCeC*	*yeCaCeC*	*CaCeC*	*leCaCeC*
Pu'al	*CuCaC*	*meCuCaC*	*yeCuCaC*	–	–
Hitpa'el	*hitCaCeC*	*mitCaCeC*	*yitCaCeC*	*hitCaCeC*	*lehitCaCeC*

Example (2) illustrates families with differential sizes—a singleton verb (*tipes* 'climb', with no other verb sharing this root), and families of two, three and five members.

1.3 Verb patterns within the *binyan* paradigm

However, a *binyan* conjugation is not isomorphic with verb pattern. This notion in fact is more complex. Each of the seven *binyan* conjugations consists of a phonologically unique bundle of five[9] temporal patterns, which combine with a root to construct the set of temporal stems—past tense, present tense, future tense, imperative[10] and infinitive. Table 1 presents the seven *binyan* conjugations as sets of temporal patterns.[11] For example, *CaCaC*, *CoCeC* and *li-CCoC* (where C's stand for root radicals) serve as the respective past, present and infinitive patterns of *Qal*. When combined with root *k-t-b* 'write', the stems *katav* 'wrote', *kotev* 'writes' and *li-xtov* 'to-write' are yielded. In the same way, patterns *hiCCiC, maCCiC, yaCCiC* and *le-haCCiC* serve as the respective past, present, future and infinitive patterns of *Hif'il*, combining with *k-t-b* to yield *hixtiv* 'dictated', *maxtiv* 'dictates', *yaxtiv* 'will dictate' and *le-haxtiv* 'to-dictate'.

Altogether, there are 31 temporal Hebrew verb patterns, forming seven paradigms, each uniquely identified with a specific *binyan* conjugation: Five non-passive *binyan* conjugations with five temporal patterns each, and two passive *binyan* conjugations with three patterns each.[12] As Table 1 shows, for most conjugations, the temporal patterns are highly similar phonologically, whereas *Qal* and *Nif'al* have more distinct temporal patterns. This organization has important implications for the acquisition and processing of the Hebrew verb lexicon, as verb stems, the most concrete

[9] Except for passive *Huf'al* and *Pu'al*, which do not have imperative and infinitive stems.

[10] In most cases, future tense and imperative forms are used for the same modal imperative purposes.

[11] The current study is derivationally oriented, and thus does not focus on the inflectional paradigm of temporal-agreement categories that each verb potentially designates. The analyses in Ashkenazi (2015) and Ashkenazi et al. (2016) pinpointed 25 inflectional categories per verb, each composed of a unique combination of tense (or mood) category plus the relevant person, number, and gender agreement categories.

[12] The nature of passive voice expression in Hebrew as derivational rather than syntactic is presented and elaborated in two recent publications (Ravid 2019; Ravid and Vered 2017).

verb forms children are exposed to and users employ, are perceived as having internal non-linear structure. Consonantal similarity with other verb stems signals a root shared in the paradigm or across a derivational family; template similarity indicates a shared verb pattern. Thus, even singleton verbs contribute to learners' perception of non-linear structure, since, like all verbs, singletons are constructed on a *binyan* paradigm and concomitantly a shared root in differently structured stems (Ashkenazi et al. 2016). For example, singleton *tipes* 'climb' is in the *Pi'el* conjugation, and accordingly root *ţ-p-s* combines with the five *Pi'el* temporal patterns to yield past tense *tipes*, present tense *metapes*, future tense *yetapes*, imperative *tapes*, and infinitive *letapes*. That is, the non-linear root and pattern structure permeates the Hebrew verb lexicon—across derivational families of root-sharing verbs with different *binyanim*, and across the set of *binyan*-specific temporal patterns within each verb.

1.4 The composition of the *binyan* system

In structural terms, the *binyan* system determines the morpho-phonological structure of verbs. This formal system is also a main vehicle serving the expression of transitivity relations. Thus, *binyan* conjugations are associated with higher or lower transitivity values, with correspondingly richer or poorer argument structures. For example, high-transitivity *Hif'il* is often associated with two or three arguments, compared with low-transitivity *Nif'al*, which mostly occurs in lower argument structures. Berman's seminal work (1993a, 1993b) was the first to show how Hebrew speaking children learn the functions of the verb system. Causativity and distinction in transitivity are lexicalized earlier than others, with high reliance on *Qal*, the *binyan* conjugation with the highest type and token frequency in Hebrew children's usage (Ashkenazi 2015; Berman and Dromi 1984; Ravid et al. 2016). The development in verb learning during the pre-school years shows that *Pi'el* and *Hif'il* mainly express high transitivity and causativity, with *Nif'al* and *Hitpa'el* mainly expressing middle voice and inchoativity, reflexivity and reciprocity (Berman and Sagi 1981; Berman 1993a, 1993b, 2003). The ability to produce reflexivity and reciprocity across the two sub-systems continues to develop during the school years, with passive verbs in *Pu'al* and *Huf'al* appearing only in late adolescence (Ben Zvi and Levie 2016; Berman and Nir-Sagiv 2007, 2010; Berman and Ravid 2009; Levie 2012; Ravid 2004; Ravid and Vered 2017).

Since they are based on verbs in different *binyan* conjugations sharing a single root, root-based derivational families combine lexically specific meanings with *Aktionsart* values such as inchoativity, causativity, reflexivity, reciprocity, middle and passive voice (see Berman and Nir-Sagiv 2004, p. 355 for a detailed table). For example, the verb family based on root *g-d-l* 'grow' consists of basic *gadal* 'grow' (*Qal*), causative *higdil* 'enlarge' (*Hif'il*), passive *hugdal* 'be enlarged' (*Huf'al*), causative *gidel* 'raise' (*Pi'el*), passive *gudal* 'be raised' (*Pu'al*), and middle-reflexive *hitgadel* 'aggrandize oneself' (*Hitpa'el*). Therefore, morpho-lexical knowledge of *binyan*-based derivational families is central in gaining command of Hebrew syntactic constructions and argument structure.

Table 2 Outline of the dual *binyan* system

Function	Older sub-system	Newer sub-system
Basic	*Qal*	*Pi'el*
Causative	*Hif'il*	*Pi'el*
Middle, Inchoative	*Nif'al, Hif'il*	*Hitpa'el*
Reflexive, Reciprocal	*Nif'al*	*Hitpa'el*
Passive	*Nif'al, Huf'al*	*Pu'al*

1.4.1 Two sub-systems

The organization of derivational verb families is linked to the internal composition of the Hebrew *binyan* system (Ravid 2019). The seven *binyan* conjugations in fact consist of two semi-redundant sub-systems (Table 2), each expressing the same set of transitivity functions and relations. What is considered to be the older sub-system—(I) *Qal, Nif'al, Hif'il*, and *Huf'al*—has most verb types and is used with most frequency (Ravid et al. 2016), while the newer system—(II) *Pi'el, Pu'al* and *Hitpa'el*—has been extremely productive since the revival of Modern Hebrew (Bolozky 2009; Schwarzwald 2002). While this classification has historical motivations (Sivan 1976), it is also currently grounded in morpho-phonological similarity (Schwarzwald 1996) and derivational affinity (Bolozky 2007): As Table 1 shows, the verb patterns in sub-system (I) manifest mostly consonant clusters (15 out of 18 verb patterns), hosting virtually only tri-consonantal roots; whereas those in sub-system (II) all have open syllables and typically host roots with three, four and even more radicals. In semantic and systemic terms, each sub-system expresses the full array of *binyan* functions, including passive counterparts for transitive conjugations (e.g., *Hif'il – Huf'al, Pi'el – Pu'al*), as well as close derivational ties among verbs within each sub-system. For example, transitive verbs in sub-system (II) very often entail their passive and middle or inchoative counterparts (e.g., *Pi'el gilgel* 'roll,Tr' – *Pu'al gulgal* 'be rolled' – *Hitpa'el hitgalgel* 'roll,Int').

This dual system is a highly efficient platform for expanding the verb lexicon across development. It enables early learning of *binyan* forms and functions and root linkage via small families of verbs within the same sub-system (Table 3), efficiently organizing lexical knowledge into categories that support the emergence of basic syntactic relations. Larger families (Table 4) express subtle differences of same-root / different-pattern combinations with similar functions across the two systems, creating semi-productive (i.e., minor or less generalizable), weak links of the type discussed in Landauer and Dumais (1997), which consolidate the *binyan* system in its lexically and morphologically rich adult form.

The linkage between *binyan* sub-systems, roots and derivational families has not been studied to date in empirical terms. Verb derivational families are taken for granted in both formal Hebrew (native and second/foreign) language instruction and psychological studies of Hebrew reading. However, little information is available regarding the actual distributions of derivational families and their components in Hebrew usage and development. A recent study on input to toddlers (Ravid et al. 2016) indicates that children are exposed to far fewer verb derivational families than pre-

Table 3 Verb families taking only one of the two sub-systems

Binyan Root	*Qal*	*Nif'al*	*Hif'il*	*Huf'al*	*Pi'el*	*Pu'al*	*Hitpa'el*
r-d-m		*nirdam*	*hirdim*	*hurdam*			
		fall asleep	put to sleep	be put to sleep			
q-n-y	*kana*	*nikna*	*hikna*	*hukna*			
	buy	be bought	cause to gain	be made to gain			
t-w-s	*tas*		*hetis*	*hutas*			
	fly		fly,Tr	be made to fly			
d-r-s	*daras*	*nidras*					
	trample	be trampled					
h-d-q					*hidek*	*hudak*	*hithadek*
					fasten	be fastened	fasten,Int
'-d-k-n					*idken*	*udkan*	*hit'adken*
					update	be updated	get updated
š-l-b					*shilev*	*shulav*	*hishtalev*
					combine	be combined	assimilate
ħ-y-k					*xiyex*		*hitxayex*
					smile		smile to oneself

Table 4 Verbs based on *binyan* conjugations in both sub-systems

binyan	*Qal*	*Nif'al*	*Hif'il*	*Huf'al*	*Pi'el*	*Pu'al*	*Hitpa'el*
g-n-b	*ganav*	*nignav*	*higniv*	*hugnav*			*hitganev*
	steal	be stolen	sneak in,Tr	be snuck in			sneak in
n-p-l	*nafal*		*hipil*	*hupal*			*hitnapel*
	fall		make fall	be made to fall			pounce
p-c-c			*hifcic*	*hufcac*	*pocec*	*pucac*	*hitpocec*
			bomb	be bombed	explode,Tr	be exploded	explode,Int
ħ-š-b	*xashav*	*nexeshav*	*hexeshiv*	*huxshav*	*xishev*	*xushav*	*hitxashev*
	think	be prominent	consider	be considered	calculate	be calculated	be considerate

viously thought, with most input consisting of singleton verbs (with no root-related verb siblings in the investigated database) and a small number of two-*binyan* families limited to one of the sub-systems. Accordingly, the main aim of the current study is to explore the emergence and consolidation of verb derivational families in the usage of children, adolescents and adults. This information will serve as the basis of a new account of the developmental origins and learning of non-linear Hebrew verb morphology.

1.5 Transparency and opacity in verb morphology

In order to acquire the Hebrew verb system, learners need to pay attention to the structural and semantic affinity of verbs sharing root skeletons in the derivational family, as well as to the similar transitivity functions among verbs sharing the same *binyan* conjugation. At a first glance, it appears that structural transparency of verb stem and semantic coherence in derivational families should sustain learning. Transparency may be compromised in one of two ways—by defective roots, creating opaque verb stems (Ravid 1990, 1995, 2012), and by low semantic coherence within the verb derivational family.

1.5.1 Structural root types

The Hebraic morphological literature classifies roots into two major formal categories—*full* and *defective* (Schwarzwald 2002). Full roots may be regarded as regular: They consist of three (or four) consonantal root radicals constructing canonical, transparent stems where root and pattern structure can be easily identified. Such verb structures, based on full roots, were illustrated in Table 1 above. Defective roots may be considered as the irregular Hebrew root category. They mostly[13] contain non-consonantal, weak radicals such as *y*, *w* or *ʔ*, yielding non-canonical, opaque stems (Berman 2003; Ravid 1995, 2012). A detailed analysis of all structural categories of roots can be found in Ravid et al. (2016) and Ashkenazi et al. (2016).

Stems based on full, regular roots are structurally transparent (Dressler 2005) in two senses—all root radicals always show up in the stem as a set of easily identifiable consonants, and the vocalic pattern of the stems is identical or similar. This is illustrated in example (3). Clearly, all verbs in the example share the same pattern *li-CCoC*, and their root radicals are transparent. Full roots thus optimize learning of the root-and-pattern non-linear structure of Hebrew verbs. In contrast, stems based on defective roots are opaque in the sense of often containing only a part of the root. For example, it is only the alternation *b*/*v* that indicates root *b-w-ʔ* in *ba* 'come' (*Qal*) and *hevi* 'bring' (*Hif'il*). Concurrently, defective roots distort the form of the stem, creating phonologically variant and fused allomorphs, which make it difficult to identify the root and pattern components and derive generalizations. This is illustrated in example (4).

(3)	*li-CCoC*	Infinitive pattern of *Qal*	
	li-shmor	'to keep' root	*š-m-r*
	li-sgor	'to close' root	*s-g-r*
	li-vdok	'to check' root	*b-d-q*
	li-gmor	'to finish' root	*g-m-r*

[13] Not all defective roots are based on glides; there are other defective categories based on consonantal roots such as the *n*-initial category (Ravid 1995). A full and detailed list is provided in Ravid et al. 2016.

(4) Infinitive *Qal* allomorphy with different defective roots

la-kum	'to-rise'	*q-w-m*
la-vo	'to-come'	*b-w-ʔ*
la-shir	'to-sing'	*š-y-r*
la-rédet	'to-go down'	*y-r-d*
li-shon	'to-sleep'	*y-š-n*
li-vkot	'to-cry'	*b-k-y*
la-cet	'to-go out'	*y-c-ʔ*

Such verbs, based on defective roots, often referring to familiar, salient activities, are highly frequent in young children's lexicons, and consequently resistant to regularizing change (Armon-Lotem and Berman 2003; Berman and Armon-Lotem 1997; Dromi 1987; Hare and Elman 1995).

A dictionary study of Hebrew roots shows that about two thirds of Hebrew root types are full (Bolozky 2007). At the same time, recent studies on input to toddlers and their output indicate that most verb tokens are based on defective roots, with full roots increasing later on. It thus appears that defective roots constitute the core verb lexicon in Hebrew, while full roots carry the burden of lexical verb learning (Ashkenazi et al. 2016:519; Ravid et al. 2016:115). In the current study, we examine this hypothesis in view of the proportions of full versus defective roots in development, and especially in the shift from spoken to written Hebrew usage.

1.5.2 Semantic (in)coherence in derivational families

The view held by many linguists and psychologists, as well as by the educated Hebrew-speaking public, is that roots carry some basic meaning shared, with semi-productive *binyan*-linked modulations, by all verbs based on the same root skeleton. This view hinges, at least in part, upon derivational verb families being indeed semantically coherent. However, a brief survey of Hebrew verbs shows different kinds of semantic relations in derivational families. Table 3 above showed small families restricted to the same *binyan* sub-system, with shared semantic senses such as sleeping, flying, updating or combining. These families also clearly linked *binyan* conjugations to transitivity relations such as causativity, passive voice and reflexive / middle voice. Table 4 showed larger, less coherent families, where lexical knowledge can still support some sort of reverse engineering. For example, the link between *ganav* 'steal' / *hitganev* 'steal in' (root *g-n-b*) can be figured out. Pouncing upon something or somebody (*hitnapel* 'pounce') may be linked to falling (*nafal* 'fall'), shared root *n-p-l*; and the whole set of mental states and activities designated by verbs based on root *ħ-š-b* makes sense. Given that verbs are lexical items, Table 4 shows derivational families that can be said to have common semantic cores in meta-linguistic thinking.

Table 5, however, depicts some less coherent phenomena. The two families at the top of the table clearly show a deep semantic shift when moving from the older sub-system to the newer one (from returning to courting, from keeping silent to being paralyzed). The next set of two families shows not only the same shift, but also ambiguity within the family, especially within the newer sub-system (agreeing and summarizing, paying and completing). The final set depicts small, restricted families that are nonetheless completely opaque semantically, so that the roots that relates

Table 5 Semantic (in)coherence in root-related verbs

binyan Root	*Qal*	*Nif'al*	*Hif'il*	*Huf'al*	*Pi'el*	*Pu'al*	*Hitpa'el*
ħ-z-r	*xazar*		*hexezir*	*huxzar*	*xizer*	*xuzar*	
	return,Int		return,Tr	be returned	court	be courted	
š-t-q	*shatak*		*hishtik*	*hushtak*	*shitek*	*shutak*	*hishtatek*
	keep silent		make silent	be made silent	paralyze	be paralyzed	quiet down
š-l-m		*nishlam*	*hishlim*	*hushlam*	*shilem*	*shulam*	*hishtalem*
		be whole	complete	be completed	pay	be paid	pay off
			make up with				cont. studies
s-k-m			*hiskim*	*huskam*	*sikem*	*sukam*	*histakem*
			agree	be agreed	agree upon	be agr'd upon	come up to
					summarize	be sum'zed	
g-z-m	*gazam*		*higzim*				
	prune		exaggerate				
ʔ-m-r	*amar*	*ne'emar*	*he'emir*				
	say	be said	increase				

pruning and exaggerating, saying and increasing respectively can only be taken as structural skeletons.[14] It thus seems that while smaller families seem to be more coherent, they can also be completely opaque, while semantic relations among verbs in larger families range from subtly modulated to opaque. In order to understand how derivational verb families based on root skeletons are learned despite structural and semantic opacity (Mattiello and Dressler 2019), we need to gain information on their distributions, size and make up across developmental corpora.

2 Research questions and hypotheses

Against this background, we posited five specific research questions and hypotheses. Our preliminary question related to the general characteristics of the study database with regards to the distributions of verbs, roots and *binyan* verb conjugations—the components of derivational verb families. We assumed that across the database, type and token usage of verbs, roots (especially full, regular roots), and *binyan* conjugations (especially low frequency conjugations) would increase and diversify with age and literacy, especially in written language. Part I of this paper focuses on these developmental changes as indicators of the growth and consolidation of the verb lexicon.

Part II starts with our next four questions and related hypotheses, focusing on the developmental characteristics of the derivational verb families in the database. First, regarding *family frequency*, we explored the numbers of verb families in the

[14]Note that in psycholinguistic priming studies, the root effect does not depend on semantic transparency—see Deutsch and Kuperman (2019).

database. Having more derivational families indicates a larger, denser and more complex verb lexicon. Therefore we assumed that the number of families would increase in discourse produced by older speakers, and especially in written language. Second, regarding *family size*, we examined the number of members in each family, i.e., the number of verb lemmas with different *binyan* conjugations sharing the same root skeleton. Larger families indicate a greater grasp of the root-and-*binyan* verb system coupled with a larger, semantically more diverse verb lexicon. Accordingly, we hypothesized an increase in family size with age and literacy. Third, we examined *family composition*, that is, the *binyan* sub-systems participating in the family. As smaller families tend to be restricted to one of the two sub-systems, we assumed that the increase in family frequency and size would be accompanied by a shift from within- to across- sub-systems. Finally, *family coherence* relates to semantic relations among derivational family members. We hypothesized that with age and literacy, verb lexicons would gradually shift from semantic transparency through polysemy to homophony.

Taken together, these hypotheses were grounded in the notion of 'starting small' (Elman 1993), that is, making use of reduced morphological entropy in first gaining command of sparse, morphological families restricted to one *binyan* sub-system, semantically coherent roots yet clearly distinguished semantic roles of *binyan* conjugations (Ackerman and Malouf 2013). A growing, morphologically and lexically diverse verb lexicon learned from variegated communicative contexts is predicted to contain more, larger and less semantically coherent derivational verb families.

2.1 Seeking distributional patterns in relevant data

As the goal of the current study is to determine the developmental route of verbs and verb families in Hebrew, the nature of the database from which this information extracted is relevant in two respects—one general and one Hebrew-specific. First, in general psycholinguistic terms, we would like to ascertain that the distributions in the database validly reflect patterns of usage that learners are exposed to (Keuleers and Marelli 2020). Frequency is one of the most important factors in adult lexical processing (Ambridge et al. 2015; Keuleers et al. 2010). When considering the role of frequency in acquisition, frequent encounters with words are, first and foremost, opportunities for lexical learning. And once a word has been acquired, greater frequency reinforces its acquisition by easing processing and serving to bootstrap the learning of other words. Importantly, and as shown by Ramscar et al. (2013), when learning new words, children's judgements about what is most informative about those words is predicted by their co-occurrence with objects and events in the environment. Taken together, to gain valid information about lexical frequencies in development, they need to come from an ecologically valid corpus, where frequencies reflect real patterns of usage that learners experience, where they truly reflect age- and modality-related changes, and where we can be fairly certain that frequencies of usage are meaningfully contextualized for learners (Behrens 2006; Goodman et al. 2008; Keuleers et al. 2015). For these goals to be met, patterns of distribution cannot be drawn from corpora consisting, for example, of Google books, movie and TV subtitles, or newspapers, as there is no evidence to what extent language learners have actually encountered them (Brysbaert et al. 2011). Word frequency in corpora has been found to correlate with age of acquisition (AoA) when said corpora consist of child-directed speech (CDS) or children's own output, child speech (CS) (Ashkenazi et al. 2016; Kidd et al. 2010; Matthews et al. 2005).

Accordingly, the database that we have amassed is composed of discourse that native Hebrew users produced and/or have been exposed to. As detailed below (the Methods section), our database contains spoken language produced by toddlers in dyadic interaction with their parents (CDS and CS) and in preschool children's peer-talk interaction. It also contains texts in two genres written by children and adolescents, as well as by younger and older adults. This means that all spoken language data is contextualized in meaningful interaction across the pre-literate years, and that the written texts are the productions of school-going populations and adults that are non-expert language users (that is, not professional or academic writers), elicited in psycholinguistically designed tasks. The only expertly written texts this database contains are popular Israeli children's storybooks that parents read to their children. Thus, the database from which the verb frequencies are extracted is ecologically valid in the senses delineated above.

But there is another, Hebrew-specific, sense in which we believe our database works well in the service of delivering accurate information about the changing patterns of verb usage in Hebrew. There are several written Hebrew corpora used for extracting lexical frequencies, some based on newspapers, others on digital resources (Frost and Plaut 2001; Itai and Wintner 2008; Linzen 2009). In addition to all of the issues elaborated above, a Hebrew specific challenge that they all fail to overcome is the extreme homography in Hebrew spelling (Bar-On et al. 2017; Ravid 2012) that permeates non-voweled Hebrew texts (i.e., virtually all texts written for native speakers aged 10 and over). This renders the computerized identification of open-class (lexical) inventories highly unreliable (Ravid et al. 2016). The only way to accurately identify each lexical item is by manually checking the string it occurs in, which is virtually impossible in these large corpora. In contrast, our database has been constructed bottom up from transcribed spoken Hebrew (see full description of the process in Ashkenazi 2015), as well as from mirror-transcribed texts written by participants in psycholinguistic studies (see full description of methodology in Berman and Verhoeven 2002). The only texts collected top-down were the children's storybooks, published in voweled script—that is, with all diacritics signifying vowels and consonantal alternations, enabling the precise identification of each word (Ravid 2012). These voweled texts were entirely analyzed by authors of this paper, ensuring full accuracy of all lexical items extracted and coded. Taken together, these two properties of our database render it a useful, reliable source of information about the emergence and growth of Hebrew verbs and derivational verb families.

3 Method

This study was conducted in a database of a total of 485,908 Hebrew word tokens, consisting of six sub-corpora, as described below. All participants contributing spoken and written discourse to the database were typically developing, monolingual, native Hebrew speakers, from middle to high SES backgrounds.

3.1 Composition of the database

(1+2). *Spoken language by toddlers and their parents*

These were two corpora of transcribed and coded child speech (72,086 word tokens) and parental child directed speech (299,461 word tokens), consisting of dense recordings (three times a week, one hour each time) over a period of six months. Participants were two Hebrew-speaking dyads (toddlers aged 1;8–2;2) in naturalistic spontaneous interaction: a boy and (mostly) his mother, a girl and (mostly) her mother (Ashkenazi 2015).

(3). *Peer talk of children aged 2–8 years*

This corpus, containing 32,991 word tokens, consisted of transcripts of conversations in six age groups of children between the ages of 2–8 years, 54 participants altogether (Eitan Stanzas 2015; Zwilling 2009). The two youngest groups of children were 2- and 2;6-year olds respectively, followed by three consecutive groups of 3-, 4- and 5-year olds, and a group of 7-year olds in 2nd grade. For each age group, three 30-minute recordings of triads of same-age children in spontaneous play were compiled to a 90-minute corpus, altogether 9 hours of transcribed and coded recordings.

(4+5). *Written text production across the school years*

This was a corpus containing 34,888 word tokens, which was compiled of two written corpora: One, 160 personal-experience narratives and ideational expositions, elicited after the screening of a video clip on the topic of "problems among people". These texts were written by 80 participants in four age groups—9–10 years (grade 4), 12–13 years (grade 7), 16–17 years (grade 11) and graduate university students (aged 25–30). These texts were a subset of a larger cross-linguistic corpus (Berman and Katzenberger 2004; Berman and Verhoeven 2002). A second corpus consisted of 300 personal-experience narratives on the topics of being offended and experiencing shame or shyness (Ravid and Hershkovitz 2017). These texts were written by 150 participants in five age groups—9–10 years (grade 4), 12–13 years (grade 7), 16–17 years (grade 11), young adults aged 19–21 during civil or military service, and university students aged 25–35.

(6). *Children's storybooks*

This corpus, containing 49,384 word tokens, consisted of children's storybooks targeted at toddlers and preschoolers, which were composed or translated by expert writers of Israeli children's literature; and school texts, primarily narratives, for beginner readers in 1st and 2nd grades (ages 6–7 years), written by child education experts (Grunwald 2014).

3.2 Coding and analyses

The current analysis focused on derivational families based on lexical verbs. The boundaries for inclusion/exclusion of forms depended on this requirement. Accordingly, the grammatical (i.e., non-lexical) root *h-y-y* 'be' was excluded from the analyses. Present-tense *beyoni* participial patterns, which are highly productive in new coinage of nouns and adjectives (Berman 1978; Ravid 2019), were included only when constituting part of the temporal paradigm of verbs. For example, adjective *mitnase* 'condescending' was excluded, while present-tense verb *mitnase* 'rising' using the same present-tense participial *Hitpa'el* pattern was included.

All tokens of lexical verbs in the corpora were identified and their derivational morphemes (root and *binyan* conjugation) coded, as elaborated below. Both type and token frequencies were analyzed where relevant, as both contribute to the emergence and entrenchment of linguistic categories in language learning and usage.

3.2.1 Verb types and tokens

Verb types were defined as verb lemmas, a unique combination of root plus *binyan* conjugation. For example, the combination of the root *b-w-ʔ* with *Qal* constituted one verb lemma (citation form, 3rd person masculine singular past tense = *ba* 'come'), while the combination of the same root with *Hif'il* constituted another verb lemma (citation form = *hevi* 'bring'). Passive verbs were counted as separate verb lemmas, given their morphological profiles in Hebrew (Ravid and Vered 2017). Verb tokens were counted as all occurrences of fully inflected verb forms (e.g., *hevénu* 'brought, 1st, Pl—we brought').

3.2.2 Root types and tokens

Root types were defined as distinct structural skeletons, so that *b-w-ʔ* 'come' was a root type distinct from, say, *b-d-d* 'separate'. Note again that roots are not verb lemmas, as the same root *b-w-ʔ* is shared by three different verbs—*ba* 'come', *hevi* 'bring', and *huva* 'be brought'. Root tokens consisted of all the occurrences of the roots in the corpus, that is, all verb tokens.

Roots were classified by their *structural categories* (Ravid et al. 2016) into full (including quadrilateral) or defective.[15] *Full* or *regular* roots are tri- and quadri-consonantal, where all root radicals appear in every inflected or derived form, yielding transparent verb structures (e.g., *higdil* 'enlarge' in *Hif'il*, based on the root *g-d-l* 'grow'; or *hit'argen* 'organize itself', based on *ʔ-r-g-n*). This category also includes roots with pharyngeal and other 'gutturals' (the so-called *groniyot* = 'made in the throat' in the Hebraic tradition; compare *hirvi'ax* 'profit', based on root *r-w-ħ*, with *higdil* 'enlarge' in the same *binyan*, based on root *g-d-l*). *Defective* or *irregular* roots are primarily those with non-consonantal radicals, including inter alia the glides *y* or *w*, the glottal *ʔ*, and the weak radical *n* which deletes in consonant clusters (Ravid 1995; Schwarzwald 2013). These defective root categories effectively change the canonical verb structure and result in opaque structures.[16]

[15] We did not use a classification of roots as containing spirantizing elements, as this was not the focus of the current analysis (Bolozky 1997).

[16] In the current analysis, classification was initially based on the traditional division of Hebrew roots by *gzarot* (structural root categories), with each instance of a root then re-examined for psycholinguistic

3.2.3 *Binyan types and tokens*

Verbs were classified by their *binyan* conjugations. All temporal verb patterns pertaining to the same *binyan* conjugation were coded accordingly. *Binyan* types constituted all verb lemma types with the same *binyan* conjugation. *Binyan* tokens consisted of all the occurrences of verb tokens with the same *binyan* conjugation in the corpus.

3.2.4 *Derivational families*

Every root skeleton served as the basis for a potential derivational verb family. The number of root-based families (both singletons and non-singleton) was calculated based on two variables—the number of verb lemmas and the number of roots, as explained below. The number of *binyan* conjugations per root type determined derivational family size. If this number was 1, the family constituted a singleton verb with no root-sharing verb relatives in the current database. This was the case, for example, of *hishta'el* 'cough', based on root *š-'-l* in *Hitpa'el*. If this number was 2, this was a two-*binyan* (or two-member) family, e.g., *samax* 'be happy' (*Qal*) and *simé'ax* 'make happy' (*Pi'el*), based on root *S-m-ħ*. Thus, family size could range from one to seven—the maximal number of *binyan* conjugations.

3.2.5 *Semantic coherence*

This was the only measure regarding derivational families that required a further quantitative analysis. To determine the semantic affinity of verbs in a derivational family, sharing the same root skeleton, we calculated the semantic relationships in pairs of root-sharing verbs. For example, regarding the four verbs in the database based on root *ħ-š-b* 'think', *xashav* 'think' was paired with *xishev* 'calculate', with *nexshav* 'be considered' and with *hitxashev* 'be considerate' respectively; *xishev* 'calculate' was paired with all the other three *ħ-š-b*-based verbs, and so forth, until all possible pairings of these four verbs (total of six) were obtained. Ambiguous verbs such as *hirkiv* 'assemble' / 'take on a ride' were paired in accordance with the number of their meanings.

This process yielded a total of 707 root-sharing pairs occurring in the database. Ten lists, each containing randomized 70–73 root-based verb pairs, were presented to 64 native-speaking experts in Hebrew developmental psycholinguistics. Each list contained a maximum of two pairs of verbs sharing the same root, placed far apart from each other. Each list was judged by 8–10 experts, who were asked to rank each verb pair by the degree of meaningful relationship on a scale of 1–5, with 1 indicating no meaningful relationship between members of the pair, and 5 indicating a strong semantic relationship. An average closeness rate for each pair was calculated, ranging

---

evidence of its current morpho-phonological behavior within the verb system. For example, root *s-b-b* 'turn around' was classified as defective (double) since, despite the transparent *Qal* form *savav* 'walk in circles', forms constructed in other *binyanim* are non-canonical—e.g., *Pi'el sovev* 'turn,Trans' compared with regular *sider* 'arrange' from regular root *s-d-r*. On the other hand, root *ħ-g-g* 'celebrate', traditionally defined as a double root, was classified here as a regular, full root, since it occurs only in transparent verb structures.

from 1–5, and the pairs were grouped into five clusters by a Model Based Latent Class Analysis (LCA) procedure.[17]

Findings are presented in two sections.[18] Results (I) (the immediately following Sect. 4) appear in Part I of this paper. Together with Sect. 5 it presents and discusses the morpho-lexical development in Hebrew verbs across our corpora in terms of the changing distributions of verbs, roots, and *binyan* conjugations across the age/literacy groups. Results (II) (Sect. 6) appear in Part II of this paper, presenting and discussing the development of root-based derivational families, including *binyan* affiliations and semantic coherence. The paper concludes in Sect. 7—the general discussion.

4 Results (I): Morpho-lexical development in Hebrew verbs

Table 6 presents the general morphological characteristics of the entire database. Table 7 presents the respective sizes of the study corpora making up this database in word and verb tokens, verb lemmas, and root types.

For the purpose of the current analyses, all corpora and sub-corpora in the database were aligned in a developmental / literacy sequence.[19] Tables 7–11 start with toddler speech production, followed by peer talk in preschool and early school age groups 2–8, written text production in schoolage populations up to adulthood, followed by spoken input to toddlers by parents, and ending in the children's storybooks written by adult experts.

Hebrew verbs offer a window on lexical growth through the changes in their components—roots and *binyan* conjugations. To capture these developmental changes across the database corpora, we examined the distributions of structural root categories, new verbs and roots, and *binyan* conjugations within respective age/literacy groups. The information presented in Tables 7–11 is followed by the interpretation and discussion Sects. 5.1–5.5 of Part I.

Table 8 shows the developmental changes in *structural root categories* across the age groups.

Table 9 depicts *the occurrence of new verb lemmas and new roots* across the database. For this purpose, the toddler speech corpus served as the baseline.

The third analysis (Tables 10 and 11) examines *the developmental changes in binyan type- and token-distributions* across the age groups.

[17] A Model Based Latent Class Analysis (LCA) enables the identification of unobservable subgroups that are similar, based on observed characteristics—in the current case, mean semantic coherence ranks for each of the verb pairs in the database (Vermunt and Magidson 2002).

[18] Note that across age groups, token counts are composed of the sum of the compiled corpora, however type counts comprise the total number of types in each corpus.

[19] We debated the place of parental input to toddlers in this developmental sequence. On the one hand, this is spoken language, while on the other hand it was produced by literate adults. Our final decision was to place this corpus based on developmental criteria (that is, with non-expert adults as the most mature group; but note that books written by experts constitute the last group), and we believe that the results indicate that this choice made sense.

Table 6 Size and composition of the entire study database

Word tokens	485,908		
Verb tokens	86,239		
Verb types	1,483		
Root tokens	86,239	Full roots	27,369 (31.7%)
		Defective roots	58,870 (68.3%)
Root types	972	Full roots	727 (74.8%)
		Defective roots	245 (25.2%)
Binyan tokens	86,239	*Qal*	59,682 (69.2%)
		Nif'al	2,662 (3%)
		Hif'il	12,490 (14.5%)
		Huf'al	60 (0.07%)
		Pi'el	8,118 (9.6%)
		Pu'al	16 (0.01%)
		Hitpa'el	3,211 (3.7%)
Binyan types	1483	*Qal*	437 (29.5%)
		Nif'al	151 (10.2%)
		Hif'il	262 (17.7%)
		Huf'al	25 (1.7%)
		Pi'el	331 (22.2%)
		Pu'al	15 (1%)
		Hitpa'el	262 (17.7%)

5 Discussion (I): Morpho-lexical development in Hebrew verbs across the learning years

To the best of our knowledge, this is the first large-scale study examining the distributions of verbs and verb components across a Hebrew database of about half-a-million word tokens, enabling the accurate identification of morphological and lexical verb properties. As over 83% of this database consisted of CDS and CS in toddlers and preschoolers' peer talk, it can be said to represent the core of the Hebrew verb lexicon. The fact that the rest of the corpora in the database, over 84,000 word tokens, consisted of written texts of adolescents and adults in two genres, as well as of texts written by experts, means that it can also provide an indication of developmental changes beyond childhood and reflect the effects of literacy. Our further analyses show that this discrepancy in text size does not hamper our ability to pinpoint later-language lexical development in verbs and roots.

Based on Tables 6 and 7, this is the information on the core distributions of Hebrew verbs found in the database: The entire database contained 86,239 verb tokens,

Table 7 Size of the study corpora in word tokens, verb tokens, verb lemmas, and root types

Age group	Word tokens	Verb tokens	Verb lemmas	Root types
Toddler Speech 1;8–2;2	72,086	7,706 (10.7%)	259	224
Peer Talk 2–2;6	2,887	531 (18.4%)	73	67
Peer Talk 2;6–3	4,012	738 (18.4%)	116	106
Peer Talk 3–4	5,165	1,026 (19.9%)	137	119
Peer Talk 4–5	5,893	1,107 (18.9%)	163	143
Peer Talk 5–6	6,712	1,215 (18.1%)	214	185
Peer Talk 7–8	8,322	1,456 (17.5%)	201	179
Peer Talk 2–8	32,991	6,073 (18.5%)	384	312
Written Text Production 9–10	3,794	894 (23.6%)	212	183
Written Text Production 13–14	7,484	1,393 (18.6%)	333	265
Written Text Production 16–17	6,678	1,262 (18.9%)	355	280
Written Text Production 19–21	3,511	750 (21.4%)	285	241
Written Text Production Adults	13,421	2,408 (17.9%)	605	458
Written Text production 9–Adults	34,888	6,707 (19.2%)	865	598
Parental Speech to Toddlers 1;8–2;2	299,461	54,810 (18.3%)	684	521
Children's Storybooks	49,384	10,943 (22.15%)	987	725

Table 8 The distribution of structural root categories across the age groups

Age group	Root tokens	Structural category		Root types	Structural category	
		Full tokens	Defective tokens		Full types	Defective types
Toddler Speech 1;8–2;2	7,706	25.4%	74.6%	224	64.3%	35.7%
Peer Talk 2–2;6	531	23.7%	76.3%	67	52.2%	47.8%
Peer Talk 2;6–3	738	22.9%	77.1%	106	64.2%	35.8%
Peer Talk 3–4	1026	31.9%	68.1%	119	63.8%	36.2%
Peer Talk 4–5	1107	27.8%	72.2%	143	64.3%	35.7%
Peer Talk 5–6	1215	36.6%	63.4%	185	66.5%	33.5%
Peer Talk 7–8	1456	33.2%	66.8%	179	71%	29%
Written Text Production 9–10	894	49.3%	50.7%	183	67.8%	32.2%
Written Text Production 13–14	1393	51.2%	48.8%	265	68.3%	31.7%
Written Text Production 16–17	1262	51.8%	48.2%	280	69.3%	30.7%
Written Text Production 19–21	750	49.7%	50.3%	241	66.8%	33.2%
Written Text Production Adults	2408	54.2%	45.8%	458	68.2%	31.9%
Parental Speech to Toddlers 1;8–2;2	54,810	27.5%	72.5%	521	73%	27%
Children's Storybooks	10,943	45.6%	54.4%	725	73%	27%

close to 18% of the word tokens. There were 1,483 different verb lemmas,[20] and 972 different roots.

[20]Note that these numbers include 140 (14%) polysemous verbs such as *hirkiv*, with two distinct meanings—'make-ride' and 'assemble'. Polysemous verbs were counted as one lemma. Their analysis requires further study.

Table 9 Lexical growth in roots and verbs across the database

Age group	New verb lemmas	New roots	Structural root category (new roots)	
Toddler Speech 1;8–2;2	259	224	Full	64.3%
BASELINE			Defective	35.7%
Peer Talk 2–2;6	7	5	Full	100%
			Defective	0
Peer Talk 2;6–3	24	15	Full	73.3%
			Defective	26.7%
Peer Talk 3–4	20	15	Full	86.7%
			Defective	13.3%
Peer Talk 4–5	29	21	Full	85.7%
			Defective	14.3%
Peer Talk 5–6	60	40	Full	70%
			Defective	30%
Peer Talk 7–8	47	38	Full	94.7%
			Defective	5.3%
Written Text Production 9–10	64	40	Full	67.5%
			Defective	32.5%
Written Text Production 13–14	119	70	Full	71.4%
			Defective	28.6%
Written Text Production 16–17	112	63	Full	84.1%
			Defective	15.9%
Written Text Production 19–21	62	40	Full	75%
			Defective	25%
Written Text Production Adults	197	113	Full	70.8%
			Defective	29.2%
Parental Speech to Toddlers 1;8–2;2	197	123	Full	87%
			Defective	13%
Children's Storybooks	286	165	Full	75.8%
			Defective	24.2%

Tables 7–11 show that the verb lexicon sampled in this study increased in size, richness and complexity with age and schooling from several perspectives. The various facets of this growth converge at a major division between the core verb lexicon, represented by the spoken discourse of toddlers with their parents and children's peer talk, and the 'advanced' verb lexicon, represented by texts written by adolescents and adults.

5.1 Verb talk in acquisition

Three initial pieces of evidence converge in outlining the emergence and growth of the verb category in Hebrew. The first is the general proportions of verb tokens as against word tokens in the database, indicating the amount of "verb talk" pro-

Table 10 The distribution of *binyan* conjugations in verb tokens

Age group	Verb tokens	*Qal*	*Nif'al*	*Hif'il*	*Huf'al*	*Pi'el*	*Pu'al*	*Hitpa'el*
Toddler Speech 1;8–2;2	7,706	80.6%	2.3%	10.7%	0%	4.7%	0%	1.75%
Peer Talk 2–2;6	531	81.9%	1.1%	7.5%	0%	7.1%	0%	2.4%
Peer Talk 2;6–3	738	75.7%	2.7%	12.3%	0%	7.2%	0%	2.1%
Peer Talk 3–4	1026	68.7%	2.1%	15.3%	0%	12%	0%	1.9%
Peer Talk 4–5	1107	71.4%	1.9%	14.4%	0%	8.8%	0%	3.4%
Peer Talk 5–6	1215	62.1%	1.7%	16.9%	0%	14.8%	0%	4.5%
Peer Talk 7–8	1456	69.5%	3%	16.3%	0%	9%	0%	2.2%
Written Text Production 9–10	894	51.4%	7.9%	18%	0.2%	14.5%	0.1%	7.9%
Written Text Production 13–14	1393	48.5%	6.5%	19.4%	0.3%	14.4%	0%	11%
Written Text Production 16–17	1262	47.8%	6.9%	17.7%	0.4%	14.9%	0.7%	11.6%
Written Text Production 19–21	750	45.2%	9.5%	20.3%	0.8%	14.1%	0%	10.1%
Written Text Production Adults	2408	41.8%	8.8%	22%	0.8%	16%	0.3%	10.3%
Parental Speech to Toddlers 1;8–2;2	54,810	71.7%	2.3%	14.5%	0%	8.8%	0%	2.7%
Children's Storybooks	10,943	62.3%	5%	13.6%	0%	12.4%	0.1%	6.6%

Table 11 The distribution of *binyan* conjugations in verb types (= verb lemmas)

Age group	Verb lemmas (types)	*Qal*	*Nif'al*	*Hif'il*	*Huf'al*	*Pi'el*	*Pu'al*	*Hitpa'el*
Toddler Speech 1;8–2;2	259	44.8%	7.7%	18.9%	0%	16.6%	0%	12%
Peer Talk 2–2;6	73	64.4%	5.5%	13.7%	0%	10.9%	0%	5.5%
Peer Talk 2;6–3	116	46.6%	8.6%	22.4%	0%	17.2%	0%	5.2%
Peer Talk 3–4	137	46%	5.6%	19.4%	0%	23.4%	0%	5.6%
Peer Talk 4–5	163	49.2%	5%	19.6%	0%	18.4%	0%	7.8%
Peer Talk 5–6	214	38.8%	6.1%	21.5%	0%	23.4%	0%	10.2%
Peer Talk 7–8	201	41.8%	7%	21.9%	0%	20.8%	0%	8.5%
Written Text Production 9–10	212	37.3%	13.2%	20.3%	0.9%	16%	0.5%	11.8%
Written Text Production 13–14	333	33%	12.9%	20.8%	1.5%	16.2%	0%	15.6%
Written Text Production 16–17	355	31%	9.3%	20.4%	2.2%	19.1%	1.4%	16.6%
Written Text Production 19–21	285	34.4%	10.2%	20.7%	2.1%	19.2%	0%	13.4%
Written Text Production Adults	605	27.8%	10.6%	21.3%	1.6%	20.2%	1.3%	17.2%
Parental Speech to Toddlers 1;8–2;2	684	36%	7.5%	18%	0.3%	21.5%	0%	16.7%
Children's Storybooks	987	36%	9.4%	17.4%	0.6%	20.2%	0.2%	16.2%

duced by participants (Table 7). Across spoken and written language, verb usage occupied over 20% of all respective corpora, appearing to be a steady property of Hebrew discourse, regardless of modality or genre. However, parent-toddler dyads stand out in this respect. Parents used verbs a little less frequently than in the general database when talking to their toddlers (18%), but the proportion of verb usage in toddlers themselves was only half as much (11%). This difference can be attributed to the fact that content-word learning, including verbs, is still very much under way in toddlers aged 1;8–2;2. To enable the expression of events, actions and states by verbs in a morphologically complex language, Hebrew-speaking toddlers need to put together root and *binyan* structure, agreement marking, and temporal and mood categories (Aguado-Orea and Pine 2015; Ashkenazi et al. 2016; Hirsh-Pasek and Golinkoff 2006; Ravid et al. 2016). This initial avoidance of verbs is made possible by another facet of Hebrew as a Semitic language, namely, the fact that verbless expression is a favored usage device (Berman 1980, 1990; Dromi and Berman 1986). This tendency is enhanced in the presence of adult caregivers, who are capable of interpreting toddlers' needs, desires and commentary despite young children's lack of discursive skills.

A second facet of verb development relates to distributions of verb roots across the study corpora (Table 8). In terms of both types and tokens, we see incremental growth in root usage within the two age-scaled lexicons of spoken peer talk (2–8 years) and written text production (9 years to adulthood). For example, 533 root tokens and 67 root types occurred in 2–2;6 year olds, in comparison to 1,456 root tokens and 179 root types in 7–8 year olds. Larger root usage indicates more verb tokens in usage, which in Hebrew includes more inflected verbforms of the same verb lemmas, more verb lemmas, and more verbs related by the same root—all pointing to a larger, denser and more diverse verb lexicon with age and literacy. The only exception to this trend are the young adults in mandatory military and civil service, whose texts contain less than half of the root tokens of 11th graders and a quarter of the tokens in the adults, and a lower number of root types as well. In addition to having a smaller word corpus (due to their being the smallest group), this may result from the fact that young adults in this database are not attending school in any form for the duration of their service. In direct contrast, the children's storybooks corpus, despite its small size, was extremely root-rich, with over 10,000 root tokens and 725 root types, by far the largest number of roots across the database, including the 55,000 verb-token parental speech corpus. Taken together, these trends indicate that literacy is a critical component in the acquisition of the Hebrew verb lexicon, promoting a wider and denser lexicon.

A third developmental perspective relates to new lexical acquisition. This was measured by the number of new verb lemmas and new roots added in each age group in comparison to those preceding it (Table 9). These increments were found to proceed in four steps. Children up to four years of age contributed 80 new verbs and 56 new roots in comparison to the baseline (Toddler Speech 1;8–2;2). Children 5–9 contributed 171 more new verbs and 118 new roots to the database. But the largest lexical enhancement came from adolescence onwards. Written texts by teenagers 13 years, 16 years and young adults added 293 new verbs and 173 new roots to the database. And adults (in both spoken and written productions) made the most contributions—680 new verbs and 401 new roots, more than all of the previous increments to the baseline together (544 verbs and 347 roots). These numbers are important, as they indicate that text size alone cannot explain the occurrence of new verbs and verb roots. Thus for example, despite the huge difference in size, the same number of new verbs (197) and a similar number of new roots (113 and 123 respectively) were added by written adults texts (13,241 word tokens, 2408 verb tokens) and by parental speech

(299,461 word tokens, 54,810 verb tokens). In our view, it is the combination of mature, proficient, densely organized adult verb knowledge, based on experience in different communicative contexts and literate expression, that makes this difference.

Beyond these developments, the derivational components of Hebrew verbs—roots and *binyan* conjugation—each deserve further, in-depth analyses to determine their contributions to verb learning across the learning years.

5.2 Structural root classes in development

Table 6 provides information on the structural composition of root types and tokens in the general database. Full (or regular, including quadrilateral) roots made up 31% of all tokens, with the rest being defective (irregular). In types, full roots made up 75% of the verb lexicon. These distributions are similar to what has been found for other languages, with irregular items having high token frequency, and regular items—high type frequency (Kuznetsova 2015; Nicoladis et al. 2007).

Table 8 shows the developmental distributions of full and defective structural classes in root types and tokens. Type-wise, across the corpora, a majority of full roots reflects the general distributions of the Hebrew verb lexicon, especially from school age onwards, with new full roots taking the lead as the major contributors of new lexical content (Table 9). Token-wise, Hebrew-speaking children are given a canonical initiation into the verb lexicon based on a small number of highly repetitive, mostly defective (irregular) roots such as *b-w-ʔ* ‘come’, so that most of the burden of lexical learning resides in the much larger repository of full roots[21] (Ashkenazi et al. 2019). But root tokens undergo a pronounced literacy shift. Spoken roots, including parental input, are overwhelmingly defective, however in the written texts, root tokens are split evenly between full and defective classes, indicating that literacy contexts accentuate the acquisition of literate verbs typically based on full roots. Interestingly, a small residue of defective roots is on the increase again in texts written by adults. For example, roots *r-’-y* ‘shepherd’ and *c-w-d* ‘hunt’ first appear in the children’s storybooks and in adults’ written productions, respectively. These defective roots are rare, literate, and lexically specific, unlike those frequently occurring in childhood.

5.3 *Binyan* distributions across development: The two sub-systems

Recall that *binyan* conjugations determine the morpho-phonological structure of Hebrew verb stems, and at the same time configure verbs into morpho-syntactic categories relating to transitivity and valence, in two *binyan* sub-systems (Sect. 1.4.1). The distributions of *binyan* conjugations in the two sub-systems (Table 6) found in the current study point to the role they play in the development of the verb system. In terms of tokens, the older *Qal-Nif’al-Hif’il* (*Huf’al*) sub-system was overwhelmingly represented (87%), as expected, with *Qal* dominating (70%) (Berman and Nir-Sagiv

[21] Note, however, that this process is typical of children from mid/high socio-economic status (SES): A new study comparing the peer talk of young children from different SES backgrounds (Levie et al. 2019) demonstrates the lesser proportion of full roots in the low SES group. This finding has implications for morpho-lexical learning gaps between the two populations.

2004, 2007; Ravid et al. 2016), while verb tokens in the newer sub-system with *Pi'el* (*Pu'al*) and *Hitpa'el* were scarce. Verb tokens in passive *binyanim* were virtually absent, as found previously (Ravid and Vered 2017). Thus, the older sub-system, especially basic *Qal*, dominates in usage, with the newer, word-churning sub-system hugely under-represented. Verb types (lemmas) presented a more balanced picture, with the older sub-system still taking the lead (60%), but *Qal* occupying less than a third of the verb types, and the newer sub-system amply represented in over 40% of verb types. Passive *binyan* conjugations still made up only 3% of the verb types. This means that children are offered the inherent transitivity relations in Hebrew (inchoativity, causativity, middle voice, reflexivity and reciprocity) through a sub-set of the *binyan* system; whereas the ability to coin new verbs crucially involves a growing familiarity with the newer sub-system, where most verbs have a much lower token frequency. These findings are supported by new analyses in Levie et al. (2019), showing that this internal organization of the *binyan* system applies even across different SES populations. The consolidation of the Hebrew verb system, critical to the construction of clause syntax, depends on the integration of the two sub-systems.

5.4 *Binyan* distributions across development: Learning to express transitivity relations

Based on Tables 10–11, four in-depth analyses of *binyan* distributions across the study corpora tell the story of learning to express transitivity relations through the *binyan* system: *Qal*, the historically core and currently most frequent *binyan* in Hebrew, with both high and low transitivity values, as in *bana* 'build' and *rac* 'run' respectively; the two high-transitivity *binyan* conjugations, *Hif'il* and *Pi'el*; the two low-transitivity *binyan* conjugations, *Nif'al* and *Hitpa'el* (Berman 1993a, 1993b); and the two passive conjugations, *Huf'al* and *Pu'al*.

5.4.1 Qal distributions

The changing distributions of *Qal* verbs contribute to the interplay between age-related and modality / literacy-related factors, and are therefore worthy of an in-depth analysis. *Qal* tokens dominated the spoken production of children (about 80%), of adult caregivers (about 70%), and to a lesser extent, also storybook texts (about 60%). Tokens steadily declined with age groups in written school texts from over 50% to the lowest *Qal* token proportion (42%) in adults' written production. *Qal* lemma types showed a similar, though steadier and more gradual decline, across the study corpora, but adult speech and children's storybooks resembled the distributions in written language (about a third of all lemmas). These findings support and complement previous analyses of input and early child speech (Ashkenazi et al. 2016; Ravid et al. 2016), showing that Hebrew-speaking children not only acquire the basic verb lexicon of Hebrew via *Qal* verbs, but also learn the foundations of verb morphology mainly through frequent encounters with and production of those core *Qal* verbs in various temporal and agreement forms.

To learn more about the role of *Qal* verbs in acquisition, we looked for those verbs that occurred in relatively large numbers (10+) in every age group, starting from toddler production. There were about 50 such verbs in seven core semantic categories: basic motion verbs such as *af* 'fly', *ba* 'come', *halax* 'walk', *nasa* 'go in car', *rac* 'run', *kafac* 'jump', *yaca* 'go out', and *zaz* 'move'; verbs denoting core events such as *avar* 'pass', *gamar* 'finish', and *kara* 'happen'; basic postures and states such as *amad* 'stand', *kam* 'get up', *nafal* 'fall', *yashan* 'sleep', and *yashav* 'sit'; general activities such as *asa* 'do', and especially those involving object manipulation such as *axal* 'eat', *laxac* 'press', *lakax* 'take', *natan* 'give', *naga* 'touch', *sam* 'put', *zarak* 'throw', *sagar* 'shut', and *patax* 'open'; core perception and mental verbs such as *ahav* 'love', *azar* 'help', *kara* 'read', *maca* 'find', *paxad* 'fear', *ra'a* 'see', *raca* 'want', *xashav* 'think', *yada* 'know', and *yaxol* 'be able'; and *dicendi* verbs such as *amar* 'say', *baxa* 'cry', *caxak* 'laugh', *ca'ak* 'scream', and *shar* 'sing'. These highly repeated verbs in child-oriented semantic classes constitute the backbone of the early Hebrew verb lexicon.

But *Qal* was also shown as the repository of lexically restricted or high-register verbs, which occurred only in the oldest age groups, and especially in written language. Some such examples are *safag* 'absorb', *xavat* 'strike', *yalad* 'give birth' (first occurrence in parental input); *arav* 'stalk, lurk', *gaval* 'border', *gazal* 'plunder', *ma'ad* 'stumble', *pacax* 'commenced', *zalag* 'leak' (first occurrence in children's storybooks); *ta'an* 'claim' (first occurrence in 13 year old texts); *asak* 'be engaged', *kalal* 'include', *marad* 'rebel, *xadar* 'infiltrate', *xal* 'be valid' (first occurrence in 11th grade texts); *xanax* 'mentor', *xashad* 'suspect' (first occurrence in young adults' texts); *maxal* 'pardon', *nazaf* 'reprove', *sata* 'go astray', *xara* 'anger' (first occurrence in adults' texts). These mostly low-frequency verbs, often *hapaxes*, clearly show that *Qal* continues to provide new labels for activities, events and states even in literate language, highlighting its centrality in the Hebrew lexicon.

Side by side with the lexical expansion in *Qal*, two sets of *binayn* conjugations link verb meaning to transitivity values.

5.4.2 Expressing high transitivity: Hif'il and Pi'el distributions

Many verbs in *Hif'il* and *Pi'el* express high-transitivity, often causative scenarios, with animate/human subjects, dynamic verbs and inanimate objects such as *hishtil* 'transplant' or *tiken* 'fix'. In terms of lemma types, almost all corpora in the database shared a similar proportion of about 40% *Hif'il* and *Pi'el* verbs, leading us to believe that these reflect the general distributions of the Hebrew verb lexicon. The developmental story is mostly told by the tokens distributions: Despite children's affinity to such scenarios, it takes them time to learn to use *Hif'il* and *Pi'el* for their expression (Berman 1993a, 1993b). Initially, high-frequency *Qal* fulfills this function, as it also does for low-transitivity scenarios. Thus, the changes in *Hif'il* and *Pi'el* token distributions across the database corpora reflect the consolidation of Semitic expression of transitivity through the *binyan* system. Tokens in these two conjugations rose from about 15% in the youngest age groups, including parental speech to toddlers and storybooks, to about 25% at the beginning of elementary school, while written texts by older children, adolescents and adults contain over 1/3 *Hif'il* and *Pi'el* verbs in usage.

As in *Qal*, there were frequently occurring *Hif'il* and *Pi'el* verb tokens from the earliest age groups, re-occurring across the database in all or most of the corpora, but they were hardly as numerous. To determine their role in acquisition, we examined to what extent these were highly-transitive, causative verbs. In *Hif'il*, there were 14 highly frequent verbs, split into two groups: one showed early alignment with the highly transitive character of *Hif'il*, including prototypical causative verbs such as *hevi* 'bring', *hixnis* 'insert', *hexlif* 'cause to exchange', *hoci* 'take out', *herim* 'take up', *horid* 'take down', *hexin* 'prepare', *hirsha* 'allow', and *her'a* 'show'. Most of these had non-causative counterparts in *Qal* and *Nif'al*, e.g., *nixnas* 'enter', *yaca* 'go out', *yarad* 'go down', and *ra'a* 'see'. A second group consisted of the highly frequent but non-causative *Hif'il* verbs *higid* 'say',[22] *hevin* 'understand', *higi'a* 'arrive', as well as two aspectual and cognitively modulating verbs (*hicl'iax* 'succeed', *hitxil* 'start'). Most new causative *Hif'il* lemmas occurred in 5 years and older groups, with lesser frequency, e.g., *heziz* 'move', *hexzir* 'bring back', *he'evir* 'move', *hifsik* 'stop', *hirtiv* 'make wet', *he'ir* 'wake up', *he'if* 'make fly', *hexbi* 'hide'. These also included cognitive and emotive verbs like *hirgi'a* 'make calm', *hirgiz* 'annoy', *he'eliv* 'insult', *hirgish* 'feel', *hexlit* 'decide', *himci* 'invent', and *hizkir* 'remind'.

In contrast, only 10 frequent, lexically basic *Pi'el* verbs occurred from the earliest to the oldest age group, none of them causative and most with lower transitivity than the frequent *Hif'il* verbs. These included *ciyer* 'paint', *diber* 'speak', *kibel* 'receive', *siper* 'tell', *siyem* 'finish', *sider* 'arrange', *sixek* 'play', *tiyel* 'stroll', *xipes* 'search', and *xika* 'wait'. Several more transitive, even causative, *Pi'el* verbs occurred with less frequency and with gaps from early childhood, including *bishel* 'cook', *cilem* 'take a picture', *mile* 'fill up', *nika* 'clean', *nigen* 'play music', *nigev* 'wipe', *perek* 'take apart', *sovev* 'turn around', *tiken* 'fix', *tipes* 'climb up', *xilek* 'distribute', *xiber* 'combine', and *xibek* 'hug'. Most of these, again, expressed basic lexical reference rather than transitivity. As in *Hif'il*, many new *Pi'el* lemmas made their first appearance around age 5 or 6 years, but unlike *Hif'il*, most of them were not causative, e.g., *biker* 'visit', *bikesh* 'ask', *gila* 'discover', *icben* 'annoy', *kimet* 'wrinkle', *kine* 'envy', *kishet* 'decorate', *litef* 'caress', *limed* 'teach', *nixesh* 'guess', *nisa* 'try', *shilem* 'pay', *shina* 'change', *tipel* 'take care of'.

This analysis reflects the specific functions of *Hif'il* and *Pi'el* in current Hebrew. *Hif'il* expresses the proto-typical function of causativity, as shown by its higher prevalence compared to *Pi'el* in discourse produced by the youngest groups, the larger number of high-frequency tokens, and their semantic content (Dattner 2015). *Pi'el*, which is more prevalent in types in the general database, is more multi-functional, with causativity only one of its functions, as indicated by the semantic analysis of types. In early child language *Pi'el* is scarcer than *Hif'il* (see also new analyses in Levie et al. 2019), while its main function as the major mechanism for new-verb derivation kicks in in the older age groups, and especially in written language, combined with the rise of full and quadrilateral root types.

5.4.3 Expressing low transitivity: Nif'al and Hitpa'el distributions

Low-transitivity scenarios are typically expressed by unaccusative, middle-voice *Nif'al* and *Hitpa'el* verbs, e.g., *nistam* 'get clogged' or *hit'alef* 'faint'. The lemma

[22] Only in the Modal Cluster (Future, Imperative and Infinitive forms), as this verb is in complementary distribution with a suppletive form (*amar* 'say') in present and past tense (Ravid 1995).

type distributions reflect the lower proportions of *Nif'al* and *Hitpa'el* in the general database (Table 6), and especially in the younger age groups (Table 11). For *Nif'al*, lemma types across the database ranged about 10% or less, while *Hitpa'el* lemmas showed an increase to over 15% in older and written production. Both *Nif'al* and *Hitpa'el* tokens were scarce across the database, except for written text production.

The qualitative analysis echoed these findings (as did the analyses in Levie et al. 2019). *Nif'al* had only two verbs that fulfilled the criteria for high frequency, that is, consistent over-10 token occurrence across all age groups: *nigmar* 'finished, all gone' and *nixnas* 'enter'. From early on, 15 more verbs occurred, albeit with fewer tokens and/or in fewer corpora. Together, these seem to make up the basic *Nif'al* lexicon, composed of the state and perception semi-auxiliary verbs *nimca* 'be there', *nir'a* 'seem', *nish'ar* 'remain', *nim'as* 'be done with', *ne'elam* 'disappear' (joined by *na'asa* 'become' and *nishma* 'sound' beyond age 6); and of the intransitive middle-voice, telic and accomplishment verbs so typical of Hebrew child language (*nidbak* 'stick', *nishbar* 'break', *nikra* 'tear', *nitka* 'be stuck' *niftax* 'open', *nishpax* 'spill', *nirdam* 'fall asleep' (joined by *nolad* 'be born' and *nisgar* 'close' beyond age 6). Two cognitive-emotive verbs (*nehena* 'enjoy', *nizhar* 'take care') occurred infrequently but early on, joined by age 6 by mental-emotive *nizkar* 'recall', *ne'elav* 'be offended', and *nirga* 'calm down'. But later-emerging *Nif'al* types showed several active, agentive verbs, such as *nicmad* 'attach oneself', or *nirsham* 'sign in', *nitla* 'hang by the hands', and *nidxaf* 'push oneself'.

Likewise, *Hitpa'el* had only one very frequent verb across all age groups—*histakel* 'look'. Early-emerging verbs with fewer tokens and/or in fewer corpora were durative or accomplishment verbs like *hishtatef* 'participate', *hitbalbel* 'become confused', *hishtana* 'change', *hitkarer* 'cool down', *hitparek* 'fall apart', *hitlaxlex* 'get dirty'; motion verbs like *histovev* 'turn around', *hitgalgel* 'roll along', and *hitgalesh* 'slide down'; and several reflexive verbs such as *hitraxec* 'wash oneself', *hitkaleax* 'take a shower', *histarek* 'comb one's own hair', *hitlabesh* 'get dressed', *hitxabe* 'hide oneself', and *hitgared* 'scratch oneself'. In addition, the early *Hitpa'el* lexicon contained basic verbs such as *hishtamesh* 'use', *hicta'er* 'be sorry' and the ubiquitous *hitkasher* 'call by phone'.

5.4.4 Passive verbs

Hebrew passive voice is expressed in two groups of *binyan* conjugations. One is the prototypically passive-dedicated *Huf'al* and *Pu'al* (e.g., *huxzak* 'be held', *xudash* 'be renewed'). Another is *Nif'al*, which expresses passive voice among other functions (e.g., *nexsax* 'be saved'). Both groups were extremely rare as verb types and in token usage. There were altogether 76 *Huf'al* and *Pu'al* tokens in the entire database, consisting of 40 types (2% and 1% respectively of all *binyan* types). *Nif'al* passives taken into consideration were only those which were unambiguously passive (e.g., *nitman* 'be buried', *nishpat* 'be judged'), excluding ambiguous *Nif'al* verbs with both passive and non-passive interpretations such as *nimca* 'be found / exist', *nirsham* 'be written / register' or *nidxaf* 'be pushed / push oneself'. They were just as rare, consisting of 35 *Nif'al* lemma types (23% of all *Nif'al* types, 2% of all *binyan* types), and 75 tokens (3% of all *Nif'al* tokens, 0% of all *binyan* tokens). All passive tokens in the database occurred only in written discourse, starting in late adolescence, and mostly appearing in written adult productions (Tables 10 and 11). These corpus analyses support the experimental results of Ravid and Vered (2017), showing that verbal passive voice is a very late developmental phenomenon in Hebrew, where several agent-demoting devices and subjectless constructions prevail (Berman 1980, 1990). Moreover, these results re-confirm Hebrew-speaking adults' preference of the two dedicated passive conjugations *Huf'al* and *Pu'al* over the multi-functional middle voice *Nif'al*.

5.5 Interim conclusion

In sum, the analysis of verb types and tokens in the database by root and *binyan* revealed two paths to Hebrew verb learning. One is the lexical path, where verbs are learned as lexical items, whose order of appearance and degree of prevalence are determined by their relevance to child language and to children's evolving experience with the world. Most of these basic verbs, often based on defective roots, were first introduced and then repeated in *Qal*, a smaller proportion by *Hif'il* and *Pi'el*, and very few by *Nif'al* and *Hitpa'el*. However, what appears as single-verb lexical learning in Hebrew has important morphological facets, given the composition of every verb by root and verb pattern.

The second, complementary, path is morpho-syntactic, relating to the transitivity values of the *binyan* conjugations, critical for the consolidation of the *binyan* system through the massive introduction of *binyan*-typical verbs, mostly with full roots, in later childhood. The causative function of *Hif'il*, and to a lesser extent, *Pi'el*, becomes apparent when highly transitive, causative verbs appear in middle childhood. Beyond sporadic innovations serving to fill lexical gaps (Berman and Sagi 1981; Ravid 1995), new-verb formation in *Pi'el* is a phenomenon that is delayed to the late school years, including denominal quadrilateral roots (e.g., *ixzev* 'disappoint', *ifsher* 'enable', and *cimcem* 'minimize'). The low-transitivity *binyan* conjugations always constitute the smallest amount of non-passive verbs, but here too, the typical properties of *Nif'al* and *Hitpa'el* become apparent only from middle childhood onwards, as before that there are not enough verb tokens to consolidate the system. Passive voice is primarily an adolescent and adult phenomenon in Hebrew.

The developmental analysis of verbs and their morphological components from toddlerhood to adulthood concludes Part I of this study of Hebrew verb acquisition. Part II below focuses on the acquisition of the system that organizes verbs into morphological root-based families from both structural and semantic viewpoints.

Part II

The goal of the current study, grounded in a new database compiled of the spoken and written productions of Hebrew-speaking toddlers, children, adolescents and adults was to offer a new, systematic account of how Hebrew root-based verb families and their components—verb lemmas, roots and *binyan* patterns—emerge and develop in structural and semantic terms, covering the long route from early childhood to adulthood.

Part I presented a general introduction to Hebrew verb morphology, the aims and hypotheses of the study, the method section, and the results and discussion sections focusing on morpho-lexical development in Hebrew verbs, roots, and *binyan* conjugations in a database containing 485,908 word tokens, 86,239 verb tokens, 1,483 verb lemmas, and 972 root types. Part II consists of the results and discussion sections regarding four facets of root-based derivational families, and the concluding discussion covering both parts.

6 Results and discussion (II): Development of Hebrew verb derivational families

In Part I above we posited five specific research questions and hypotheses. The first question related to the general characteristics of the study database with regards to the developmental distributions of verbs, roots and *binyan* verb conjugations—the components of derivational verb families. This analysis of the growth and consolidation of the verb lexicon revealed two parallel paths to verb learning in Hebrew—the lexical path, where verbs are learned as lexical (though morphologically-oriented) items; and the morpho-syntactic path, relating to the transitivity values of the *binyan* conjugations.

We now turn to the analysis of derivational families in developmental perspective, focusing on the remaining four research questions, all pertaining to derivational root-based families. We present results and discussion in four distinct sections, each corresponding to a research question: *family frequency* (the number of root-related families in the database), followed by *family size* (the number of members in each family), *family composition* (the *binyan* make-up of families), and finally *family coherence* (the degree of semantic relatedness between pairs of root-related verbs). As these are all essentially measures of lexical density and diversity, the current derivational family analyses mostly focused on verb lemmas.

6.1 Family frequency

The notion of 'family frequency' was examined in two ways. First, the frequency of co-occurrence of root-sharing verbs in the full database, reflecting the number of derivational families in this sample of spoken and written Hebrew. Second, the frequency of co-occurrence of root-sharing verbs in each of the corpora making up this database, reflecting the number of derivational families produced and experienced in the discourse of participants in a certain age group.

The total number of different roots that occurred in the database was 972. In this list, 588 (60%) roots were singletons, that is, they occurred in one *binyan* only, without demonstrating other family members in the current database (see similar results in Ashkenazi et al. 2019 and Levie et al. 2019). For example, *sheret* 'serve' in *Pi'el*, or *nish'an* 'lean' in *Nif'al*. The rest of the roots (384, 40%) had derivational verb families, that is, they occurred in more than one verb, and concomitantly with more than one *binyan*. For example, root *q-p-c* occurred in the database in a three-verb family—*Qal kafac* 'jump', *Hif'il hikpic* 'make jump', and *Pi'el kipec* 'hop'; and root *h-r-s* occurred in a two-verb family—*Qal haras* 'destroy' and *Nif'al neheras* 'get destroyed'.

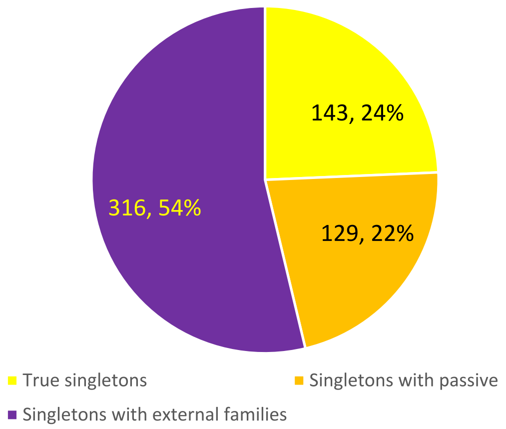


Fig. 1 Status of singletons in the database ($N = 588$) (Color figure online)

6.1.1 Singletons in the database

Why were singletons the largest group of roots in the database, and what implications does this pattern have for learning Hebrew verbs in morphological families? A first step towards answering these questions was determining whether they were true singletons in Hebrew, or whether they had "siblings" (i.e., root-sharing verbs) which did not show up in our database. To this end, every singleton verb was scrutinized for possible family members external to the database. Figure 1 shows that the 588 singleton verbs in our database roughly fell into two groups. Close to one half (46%, 272 verbs) were indeed singletons. This group in fact consisted of two categories—true singletons (24%, 143 verbs) with no other verb sharing their root (e.g., *sha'ag* 'roar' or *hishta'el* 'cough'); and singletons whose only other family member was an external, dedicated passive form: e.g., *hish'a* 'suspend' with a dedicated external passive *hush'a* 'be suspended' (22%, 129 verbs). In the latter case, it was almost always the active member that occurred in the database—only 3 were passive verbs that had an external active counterpart—e.g., *huxtam* 'be stained', with external *hixtim* 'stain'. Once more, this finding testifies to the rarity and high register of true passives in Hebrew (Ravid and Vered 2017).

Over one half of the singletons in the database (54%, 316 verbs) were not true singletons; rather, they had external families whose distributions are depicted in Fig. 2. This figure shows that the most numerous singletons constituted part of families with three members—most typically containing a transitive, a passive and a middle-voice member (e.g., *mimesh* 'realize [make true]', with external *mumash* 'be realized' and *hitmamesh* 'became true'). Two large groups were singletons in two-member (but non-passive) families (e.g., *tama* 'wonder', with external *hitmí'a* 'make wonder') and in four-member families (e.g., *ratax* 'boil', with external *hirtí'ax* 'make boil', *hurtax* 'be boiled', and *hitrate'ax* 'erupt in anger'). Larger families were more scarce, but nonetheless there were 33 singletons belonging to families with 5–7 members, e.g., *mazag* 'pour' with external *nimzag* 'be poured', *mizeg* 'blend,Tr', *muzag* 'be blended', and *hitmazeg* 'blendin,Int'. While this analysis is restricted to the singletons in our database, it highlights the intriguingly central role of singletons in Hebrew discourse, and by extrapolation, the distributions of derivational verb families in general.

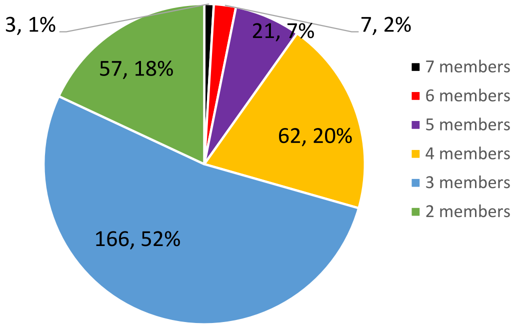


Fig. 2 Distribution of singletons with external families ($N = 316$) (Color figure online)

As each age group is represented by a different corpus, Table 12 presents the distributions of singleton roots and roots relating derivational families for each group, as well as in the entire database.

6.1.2 Singletons in verb development

A closer look at the family frequencies within each of the age groups shows that singletons dominate children's (mainly) spoken productions, constituting 85–90% of the verb roots up to age 10. But even in written productions by adolescents, there were about 80% singletons on the average, with the lowest proportions (over 70%) in adults' CDS, written texts, and in expert-written children's storybooks. Thus, despite the vast difference in text size and verb tokens (13,421 words and 2,408 verbs in adults' written texts, 299,461 words and 54,810 verbs in adults' CDS, and 49,384 words and 10,943 verbs in children's storybooks), all adult language productions consistently contained more verb families than those of adolescents and children. As more derivational verb families indicate a larger and denser verb lexicon, this developmental picture once again reflects the critical role of age and schooling in language acquisition.

As shown in Figs. 1 and 2, one reason there are so many singleton verbs in our database is because one third of all verbs (one half of the singletons) are indeed singletons in Hebrew.[23] Note, however, that being a singleton verb does not imply the absence of verb morphology nor the absence of non-verb derivational siblings. For example, *hit'akesh* 'act stubbornly' is a singleton verb in *Hitpa'el*, related by root to two adjectives meaning 'stubborn'—Biblical *ikesh* and current *akshan*. But a general reason underlies the limitation on the usage of related verbs at the developmental interface between lexicon and derivation. As verbs are recruited for the expression of events and states as lexical items serving communicative purposes, there are few opportunities for the inclusion of same-family members in the discourse. For example, see singletons *histakel* 'look' in toddler production, or *bagad* 'betray' in the young adults group. A co-occurrence of verbs from the same derivational family implies the expression of a transitivity shift, as in *ha-balon hitpocec* 'the-balloon blew up

[23]In the spoken language addressed to and produced by children, the possible passive counterparts are irrelevant (Ravid and Vered 2017).

(exploded)' / *aba pocec oto* 'daddy blew it up'. This requires a combination of a communicative opportunity, two clearly related verbs, and the derivational ability to express this relationship. Rich communicative contexts and densely lexical language usage, which are most usage-friendly for expressing overt derivational ties, are linked to age and literacy. These predictions are taken up in the next sections on family size, composition and semantic coherence.

To sum up this section, a majority of the verbs Hebrew speakers (and writers)—especially young children—experience and produce are not packaged in the traditionally conceived derivational family format but rather as discrete (yet morphologically complex) lexical items. The structures they share facilitate the emergence of the systematic Semitic notions of root, *binyan* and derivational family, with most derivationally connected forms provided in literate, written language later on.

6.2 Family size

The notion of 'family size' relates to the co-occurrence of verbs related by the same root, regarding the number of such different verbs in different *binyan* conjugations. We are interested in the size of derivational families in the database, as a representative sample of spoken and written Hebrew across the learning years and in adult usage; and also in the typical sizes of families in each of the age-defined corpora making up the database. Table 12 presents this information. It shows that derivational verb family size increased as predicted across the age-related corpora making up the database in two senses. One, in the sense of consisting of fewer singletons and more root-related families, especially in the older age groups (teenagers and above); and two, in the sense of families growing larger, from virtually only two-member families in the younger groups to three- and four-family members in the older groups, and most especially in written language. It is important to note that, as with the singletons, many of these families must be partial, with external members that do not show up in our database. But as with the singletons, this analysis can be said to typically reflect the distributions and properties of derivational verb families produced and encountered by non-expert Hebrew users.

6.2.1 The default two-member derivational verb family

A comparison of derivational family sizes across the database groups makes it clear that the most prevalent family size both in the database as a whole (over a quarter of all root types) and in each of the age groups is a two-member family. Two-member families thus seem to constitute the default derivational verb family in Hebrew usage—that is, the most frequent co-occurrence of root-related verbs in discourse production is restricted to two. The discussion below, followed by the discussion of families in terms of *binyan* composition and semantic coherence, sheds a new light on how Hebrew derivational verb families are learned from usage.

In terms of size, up to age 9, it was mostly two-member families (on the average about 10%) that complemented the overwhelmingly singleton-verb lexicon in use. In these young and (mostly) spoken productions, the co-occurrence of same-family verbs always took place based on high frequency roots and always implemented the

Table 12 The distribution of derivational families within each age group

Age group	Roots (types)	Singletons	2 members	3 members	4 members	5 members	6–7 members
Toddler Speech 1;8–2;2	224	193 (86.2%)	28 (12.4%)	2 (1%)	1 (0.4%)	0	0
Peer Talk 2–2;6	67	62 (92.5%)	4 (6%)	1 (1.5%)	0	0	0
Peer Talk 2;6–3	106	96 (90.5%)	10 (9.5%)	0	0	0	0
Peer Talk 3–4	119	102 (85.8%)	16 (13.4%)	1 (0.8%)	0	0	0
Peer Talk 4–5	143	124 (87%)	18 (12.5%)	1 (0.5%)	0	0	0
Peer Talk 5–6	185	158 (85.4%)	25 (13.5%)	2 (1.1%)	0	0	0
Peer Talk 7–8	179	158 (88.2%)	20 (11.2%)	1 (0.6%)	0	0	0
Written Texts 9–10	183	155 (84.7%)	27 (14.8%)	1 (0.5%)	0	0	0
Written Texts 13–14	265	206 (77.8%)	51 (19.2%)	7 (2.6%)	1 (0.4%)	0	0
Written Texts 16–17	280	213 (76%)	59 (21.1%)	8 (2.9%)	0	0	0
Written Texts 19–21	241	202 (83.8%)	34 (14.1%)	5 (2.1%)	0	0	0
Written Texts Adults	458	340 (74.2%)	95 (20.8%)	18 (4%)	4 (1%)	1	0
Parental Speech to Toddlers 1;8–2;2	521	375 (72%)	130 (25%)	15 (2.8%)	1 (0.2%)	0	0
Children's Storybooks	734	515 (70.2%)	169 (23%)	47 (6.5%)	3 (0.4%)	0	0
Entire Database	972	588 (60.4%)	273 (28.1%)	89 (9.2%)	20 (2.1%)	2 (0.2%)	0

most typical *binyan* transitivity modulations. For example, toddlers' productions contained transitive *shavar* 'break' with middle-voice, intransitive *nishbar* 'break' (root *š-b-r*), ergative *nixnas* 'enter' with causative *hixnis* 'insert, bring in' (root *k-n-s*), and transitive *lixlex* 'dirty,Tr' with middle-voice *hitlaxlex* 'get dirty' (root *l-k-l-k*). In the same way, the 2–2;6 peer talk had ergative *yaca* 'go out' and causative *hoci* 'take out' (root *y-c-ʔ*), and the 2;6–3 peer talk contained *paxad* 'fear' together with causative *hifxid* 'frighten' (root *p-ħ-d*). The 3–4 peer talk contained transitive *axal* 'eat' and causative *he'exil* 'feed' (root *ʔ-k-l*), as well as transitive *iper* 'make up' and reflexive *hit'aper* 'make oneself up' (root *ʔ-p-r*). The 4–5 peer talk had transitive *patax* 'open' and telic, middle-voice *niftax* 'open up'. The 5–6 peer talk contained middle-voice *nidbak* 'stick,Intr' and causative *hidbik* 'glue,Tr' (root *d-b-q*); ergative *zaz* 'move' and causative *heziz* 'move, Tr' (root *z-w-z*); as well as canonical *nish'ar* 'remain, stay' with transitive *hish'ir* (root *š-ʔ-r*). And the 7–8 peer talk contained canonical, middle-voice *ne'evad* 'get lost' together with more agentive *ibed* 'lose'; *yada* 'know' and causative *hodí'a* 'announce' (root *y-d-'*); transitive *kilef* 'peel' and middle-voice *hitkalef* 'peel off' (root *q-l-p*).

These co-occurring, root-related two-member families all expressed canonical valence-changing perspectives on prominent scenarios in children, mainly having to do with existence, possession, and motion events, with a few canonical cognitive verbs. This is further illustrated by the three two-member families prevalent across all age groups: basic *ba* 'come'—causative *hevi* 'bring' (root *b-w-ʔ*); transitive *gamar* 'finish'—middle-voice, telic *nigmar* 'all gone' (root *g-m-r*); and ergative *yarad* 'go down'—causative *horid* 'take down' (root *y-r-d*).

In older, mostly written age group productions, family size increased quantitatively, with two-member families now occupying 20% and over of the roots in each corpus. But size also changed qualitatively. On the one hand, many of the two member families in these corpora continued to demonstrate transitivity-modulating relationships in frequent verbs denoting motion, existence and presentative events, reflected in the fact that ergative *xazar* 'come back' with causative *hexzir* 'return,Tr' co-occurred in all age groups five years and older (root *ħ-z-r*). Among other typically occurring verbs in older age groups there were inchoative *hitmale* 'fill up' and causative *mile* 'fill up,Tr' (root *m-l-ʔ*) in the corpus of parental talk to toddlers, and *higí'a* 'arrive' and *naga* 'touch' (root *n-g-'*) in the 11 grade written texts. On the other hand, two-member families in the older groups also involved lexically rarer and less canonical verbs. For example, *avar* 'pass' and causative *he'evir* 'make pass', *parax* 'blossom' and causative *hifrí'ax* 'make blossom', *hirgish* 'feel' and middle-voice *hitragesh* 'get excited', in the children's storybooks; *kam* 'get up' and causative *hekim* 'raise', *risek* 'crush' and middle-voice *hitrasek* 'crash' in the parental input; cognitive *lamad* 'learn' with its causative counterpart *limed* 'teach' (root *l-m-d*) in the written 9–10 year old texts; telic *parac* 'burst' with middle-voice durative-accomplishment *hitparec* 'burst out' in 13–14 year olds' written texts; and *sider* 'arrange' with *histader* 'arrange oneself' (root *s-d-r*) in the young adults. Also, the older groups, especially in written discourse, produced co-occurring pairs of active/passive verbs, e.g., *bikesh* 'ask' and *hitbakesh* 'be asked', *natan* 'give' and *nitan* 'be given', *hisig* 'gain' and *husag* 'be gained' in the children's storybooks; *hefic* 'distribute' and *hufac* 'be distributed' in the 9–10 year olds' written texts; *hevix* 'embarrass' and *huvax* 'be embarrassed', *hicig* 'present' and *hucag* 'be presented' in the 13–14 year olds' texts; *tixnen* 'plan' and *tuxnan* 'be planned' in the 16–17 year olds' texts; *hizmin* 'invite' and *huzman* 'be invited', *shalax* 'send' and *nishlax* 'be sent' in the young adults; and *patar* 'solve' and *niftar* 'be solved', *kava* 'determine' and *nikba* 'be determined', *te'er* 'describe' and *to'ar* 'be described', *hidgish* 'emphasize' and *hudgash* 'be emphasized' in the adults' written texts. As we showed above, and elsewhere (Ravid and Vered 2017), the production of genuinely (that is, non-adjectival) passive verbs is a hallmark of literate, abstract and detached mature Hebrew usage.

6.2.2 Larger families

The co-occurrence of three or more root-related verbs in derivational families was extremely restricted. The single 3-member family that occurs in almost all of the corpora produced by children aged 2–10 is based on root *r-ʔ-y* 'see', with *ra'a* 'see', *her'a* 'show', and the frequent perception expression *nir'a li* '(it) seems to me'. It is only in the discourse produced in written texts of adolescents that several different 3-member families finally appear, amounting to 4% of the verb roots in each of the corpora. For example, *patax* 'open,Tr', *niftax* 'open,Int', *hitpaté'ax* 'develop,Int' (root *p-t-ħ*), and *pana* 'turn', *hifna* 'refer', *hufna* 'be referred' (root *p-n-y*) in the 13–14 year olds; *ne'elam* 'disappear', *he'elim* 'make disappear', *hit'alem* 'ignore' (root *'-l-m*); *maca* 'find', *nimca* 'be found, exist', *himci* 'invent' (root *m-c-ʔ*), and *nitpal* 'pick on', *tipel* 'take care of', *tupal* 'be taken care of' (root *ṭ-p-l*) in the 16–17 year olds; *yaca* 'leave', *hoci* 'take out', *yice* 'export' (root *y-c-ʔ*), *nigash* 'approach', *higish* 'bring near, serve', *hugash* 'be served' (root *n-g-š*), and *xashav* 'think', *hexhshiv* 'consider', *hitxashev* 'be considerate' (root *ħ-š-b*) in the young adults aged 19–21. All of these examples show members with derived rather than transitivity-modulated lexical semantics, such as *hitpaté'ax* 'develop,Int', *hit'alem* 'ignore', *himci* 'invent', or *yice* 'export', which demonstrate systematic morphological knowledge about the notion of derivational verb family supported by a broader variety of topics, themes and communicative functions in these written narrative and expository texts.

The corpus of written adult texts (2,408 verb tokens in 13,421 word tokens) and the vastly larger parental speech corpus (54,810 verb tokens in 299,461 word tokens) yielded a similar number and proportion of three-member derivational verb families (18, 4% in written texts, 15, 3% in parental speech)—showing, again, the immense impact of literacy on morphological family size. Side by side with extended transitivity modulations (*yashav* 'sit', *hoshiv* 'seat', *hityashev* 'sit down', root *y-š-b*; and *af* 'fly', *he'if* 'make fly', *hit'ofef* 'fly away', root *'-w-p*), these families involved high-register verbs with looser semantic ties, as in *hiskim* 'agree', *sikem* 'conclude/summarize', *histakem* 'amount to' (root *s-k-m*), or *raca* 'want', *rica* 'placate', *hitraca* 'consent' (root *r-c-y*). As predicted, the largest family size distributions were found in the corpus of children's storybooks (10,943 verb tokens in 49,384 word tokens), with 7% (47) 3-member families with extended modulations (e.g., *dalak* 'burn', *nidlak* 'turn on', *hidlik* 'turn on,Tr', root *d-l-q*; *lavash* 'wear', *hilbish* 'dress,Tr', *hitlabesh* 'dress up', root *l-b-š*), and depicting high-register and lexically specific verbs such as *nicav* 'stand still', *hiciv* 'set up', *hityacev* 'present oneself' (root *y-c-b*); and *hispik* 'suffice', *sipek* 'supply', *histapek* 'settle for' (root *s-p-q*).

The ceiling family size in larger, written, adult-produced corpora was thus three, with larger families absent in virtually all separate corpora. Only two corpora—written texts by adults, and children's storybooks—had a few 4-member families. Two examples are *shana* 'peruse', *nishna* 'repeat,Int', *shina* 'change', *hishtana* 'change,Int' (root *š-n-y*) in the storybooks; and *kadam* 'precede', *hikdim* ,bring forward', *kidem* 'promote', *hitkadem* 'make progress' (root *q-d*-m) in the written adult texts. It was only in the full, compiled database—a sample of native Hebrew usage—that larger families were represented (Table 12). In the full database, about 40% of the roots showed families, with over 1/4 of the roots participating in two-member families, and about 12% in larger families of 3, 4 and even 5 members.

6.2.3 Type and token distributions within a family

So far, derivational families were portrayed with the verb lemma lexicon in mind, with a single occurrence of a verb lemma sufficing to be counted as a member of a derivational family. Thus, the two-*binyan* family based on root *t-p-s* in the 4–5 peer talk production consisted of one token of each lemma—one of *tafas* 'catch' in *Qal* and one of *nitpas* 'get caught' in *Nif'al*. But this was definitely not the case across the board, as the number of token occurrences of each family member was not usually balanced. The analysis of derivational family size cannot be concluded without briefly attending to the issue of verb token frequency in derivational families.

Each verb lemma in Hebrew consists of a paradigm of 25–28 wordforms, each expressing a unique designation of temporal category with the relevant agreement markers of person, number and gender (see full depiction in Ashkenazi 2015 and Ashkenazi et al. 2016). Token frequency of a single verb lemma in a corpus can thus be attributed to occurrences of several wordforms from the paradigm (e.g., *nafalti* 'I-fell' and *yiplu* 'they-will-fall'), as well as to re-occurrences of each wordform. Repetition of the same wordform is used to highlight and entrench a temporal / agreement configuration in discourse, while usage of different wordforms from the same paradigm can be treated as cases of resonance, serving to comment on, maintain and expand discourse topics. It is thus clear that an analysis of the semantics and discursive context of verb token frequencies can inform us about the ways the semantic space of a verb is cognitively engaged in a text, as well as about the reasons one member of a derivational family is so highly prevalent. While the scope of the current study cannot permit an in-depth analysis, some observations about the general patterning of token distributions in families (disregarding wordform type) are called for.

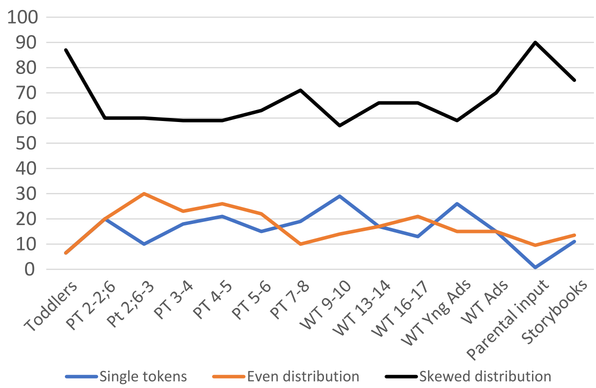
Table 13 and Fig. 3 provide information on the token frequencies of members in verb derivational families. They depict three rough categories of token patterning in these families—single, balanced, and skewed. One pattern consisted of each member of the derivational verb family being represented by a *single* token, as in the example above, or in the family based on root *p-n-y: pana* (*Qal*) 'turn', *hifna* (*Hifi'l*) 'direct, refer to', and passive *hufna* (*Huf'al*) 'be directed, referred to' (written texts, 13–14 year olds). This pattern had an especially low distribution in parental input to toddlers and in the toddlers' own speech, and occupied between 10–20% of all families in most other age groups. A second pattern involved each member of the verb family having the same number of tokens. Without exception, *balanced* verb families of this type were small, consisting of 2–5 members. For example, the family based on root *b-ħ-n* had 3 tokens for *baxan* (*Qal*) 'examine' and 3 for *hivxin* (*Hif'il*) 'observe' (children's storybooks); and in the 16–17 year olds' written texts, three verbs based on root *p-t-ħ* each had 4 tokens—*patax* (*Qal*) 'open, Tr', *niftax* (*Nif'al*) 'open,Int', and *hitpate'ax* (*Hitpa'el*) 'develop,Int'.

The overwhelming majority (60–90%) of derivational families, however, displayed the third, *skewed* pattern (Fig. 3), where one family member heavily outnumbers the others in tokens. Table 14 provides examples of such families across all the age groups. In larger corpora (such as the spoken parent-toddler interaction) and in lexically denser corpora (such as the adult written texts or the children's storybooks)

Table 13 Derivational family categories in terms of token distributions

Age group	# Families	Single tokens only	Balanced token distributions	Skewed token distributions
Toddler Speech 1;8–2;2	31	2 (6.5%)	2 (6.5%)	27 (87%)
Peer Talk 2–2;6	5	1 (20%)	1 (20%)	3 (60%)
Peer Talk 2;6–3	10	1 (10%)	3 (30%)	6 (60%)
Peer Talk 3–4	17	3 (18%)	4 (23.5%)	10 (59%)
Peer Talk 4–5	19	4 (21%)	5 (26%)	10 (59%)
Peer Talk 5–6	26	4 (15%)	6 (22%)	17 (63%)
Peer Talk 7–8	21	4 (19%)	2 (10%)	15 (71%)
Written Texts 9–10	28	8 (29%)	4 (14%)	16 (57%)
Written Texts 13–14	59	10 (17%)	10 (17%)	39 (66%)
Written Texts 16–17	67	9 (13%)	14 (21%)	44 (66%)
Written Texts 19–21	39	10 (26%)	6 (15%)	23 (59%)
Written Texts Adults	118	18 (15%)	18 (15%)	82 (70%)
Parental Speech to Toddlers 1;8–2;2	146	1 (0.7%)	14 (9.5%)	131 (90%)
Children's Storybooks	219	25 (11%)	30 (13.5%)	168 (75%)
Entire Database	384	28 (7.1%)	90 (23.5%)	266 (69.4%)

Fig. 3 Token distributions (percentages) in derivational families (Color figure online)



there were family members with hundreds of tokens, whereas in smaller or less lexically dense corpora the token-numerous members were less numerous, but the same pattern occurred across all corpora. The highest prevalence of skewed families was found in toddlers' speech and parental input to them, where repetition and resonance of specific verbs and verbforms serve the communicative context of teaching and learning.

Table 14 provides some examples of skewed distributions in 2- and 3-member families. The overall pattern in all of these examples is the high frequency of the most basic or default member of the verb derivational family, in contrast to a very low frequency of (a) more marked member(s). One pattern shown here is a high-frequency,

Table 14 Examples of derivational families with skewed token distributions

Age group	Root	Member with numerous tokens	Member with few tokens	Member with few tokens
Toddler Speech 1;8–2;2	*š-m-‘*	10 *shama* ‘hear’, *Qal*	1 *nishma* ‘sound’, *Nif’al*	1 *hishmi’a* ‘make hear’, *Hif’il*
Peer Talk 2–2;6	*r-ʔ-y*	50 *ra’a* ‘see’, *Qal*	1 *nir’a* ‘seem’, *Nif’al*	1 *her’a* ‘show’, *Hif’il*
Peer Talk 2;6–3	*y-r-d*	5 *yarad* ‘go down’, *Qal*	1 *horid* ‘take down’, *Hif’il*	–
Peer Talk 3–4	*s-g-r*	23 *sagar* ‘close’, *Qal*	2 *nisgar* ‘close down’, *Nif’al*	–
Peer Talk 4–5	*š-ʔ-r*	9 *nish’ar* ‘remain, *Nif’al*	3 *hish’ir* ‘leave’, *Hif’il*	–
Peer Talk 5–6	*ʔ-r-g-n*	19 *irgen* ‘organize’, *Pi’el*	1 *hit’argen* ‘get organized’, *Hitpa’el*	–
Peer Talk 7–8	*q-š-r*	6 *hitkasher* ‘call up’, *Hitpa’el*	1 *kashar* ‘tie’, *Qal*	–
Written Texts 9–10	*b-y-š*	39 *hitbayesh* ‘feel shy/ashamed’, *Hitpa’el*	1 *biyesh* ‘shame’, *Pi’el*	–
Written Texts 13–14	*l-m-d*	45 *limed* ‘teach’, *Pi’el*	1 *lamad* ‘learn’, *Qal*	–
Written Texts 16–17	*ħ-š-b*	24 *xashav* ‘think’, *Qal*	1 *hexshiv* ‘consider’, *Hif’il*	1 *hitxashev* ‘be considerate’, *Hitpa’el*
Written Texts 19–21	*y-c-ʔ*	11 *yaca* ‘go out’, *Qal*	3 *hoci* ‘take out’, *Hif’il*	1 *yice* ‘export’, *Pi’el*
Written Texts Adults	*k-n-s*	27 *nixnas* ‘enter’, *Nif’al*	1 *hixnis* ‘insert’, *Hif’il*	2 *kines* ‘gather,Tr’, *Pi’el*
Parental Speech to Toddlers 1;8–2;2	*k-w-n*	161 *hexin* ‘prepare’, *Hif’il*	10 *hitkonen* ‘prepare oneself’, *Hitpa’el*	–
Children’s Storybooks	*y-š-b*	206 *yashav* ‘sit’, *Qal*	1 *hoshiv* ‘seat’, *Hif’il*	24 *hityashev* ‘settle down’, *Hitpa’el*

lower agency (*xashav* 'think', *hitbayesh* 'feel shy') or low transitivity (*yarad* 'go down', *yaca* 'go out') verb, mostly in *Qal*, but also in *Nif'al* and *Hitpa'el*. This usage-prominent member is countered by a low-frequency, higher agency, causative verb in *Hif'il* or *Pi'el* respectively (*her'a* 'show', *hexshiv* 'consider', *biyesh* 'shame', *horid* 'take down', *hoci* 'take out'). An opposite pattern has a highly agentive, transitive verb (*irgen* 'organize', *limed* 'teach', *hexin* 'prepare') in *Pi'el* or *Hif'il* as the high-frequency member, with a lower-agency low-frequency counterpart (*hit'argen* 'get organized', *lamad* 'learn', *hitkonen* 'prepare oneself') in *Hitpa'el* or *Qal*.

While this phenomenon deserves a separate, statistically oriented investigation, as well as a deeper qualitative classification, this preliminary analysis can shed some light on the process of Hebrew verb learning. First, it is clear that the notion of 'default' and 'marked' member depends on the morpho-syntactic configurations provided by the *binyan* system, verb and root semantics, and pragmatic-cognitive factors such as discourse type and age group. Secondly, it questions the notion of 'acquisition' prevalent mostly in the generativistic community, where a single appearance of a form is considered as a sign of its being 'acquired'. Rather, it seems that the high-frequency, cognitively prominent family member serves as a platform for more semantically complex, less prominent family members sharing the same root.

6.3 Interim conclusion: Family frequency and size

What we show here is that the meaning and structure of Hebrew verbs is not learned directly from co-occurrences of multiple verbs in different *binyan* conjugations sharing the same root. It starts by relating clusters of temporal patterns into coherent *binyan* conjugations, which helps children construe the basic transitivity values of the *binyan* system, despite the scarcity of root-related families in their productions and in the input they experience (Ravid et al. 2016). Our findings show that the *binyan* system and, by necessity, derivational families, start small with two-*binyan* families consistently expressing core transitivity contrasts—basic vs. causative, basic vs. inchoative, middle vs. transitive or causative, in verbs with shared root skeletons. The abstract notion of verb derivational family emerges from long and extensive experience with the system in a variety of communicative contexts, together with developmental changes in cognitive, linguistic and literacy abilities over the learning years. The next two sections highlight further morphological and semantic facets of this path into learning the Hebrew verb system.

6.4 Family composition

We have already seen that most verbs encountered in the native, non-expert production of speakers and writers are not engaged in derivational families, and that most verbs that do co-occur with other verbs sharing the same root constitute small derivational families. The question at hand is what determines *binyan* distributions within a derivational verb family. To this end, we have defined the notion of *family composition* as referring to the internal distribution of *binyan* conjugations within the two sub-systems. Recall that the *binyan* system described in the introduction (Sect. 1.4) falls into two sub-systems—the older system of *Qal-Nif'al-Hif'il-Huf'al*,

Table 15 The sub-system composition of derivational families in the database corpora

Age group	# Families	A single sub-system		Both sub-systems
		Old	New	
Toddler Speech 1;8–2;2	31	22 (71%)	5 (16%)	4 (13%)
Peer Talk 2–2;6	5	5 (100%)	0	0
Peer Talk 2;6–3	10	9 (90%)	0	1 (10%)
Peer Talk 3–4	17	14 (82.4%)	1 (5.8%)	2 (11.8%)
Peer Talk 4–5	19	13 (68.4%)	1 (5.3%)	5 (26.3%)
Peer Talk 5–6	26	14 (53.8%)	8 (30.8%)	4 (15.4%)
Peer Talk 7–8	21	15 (71.4%)	3 (14.3%)	3 (14.3%)
Written Texts 9–10	28	21 (75%)	5 (17.8%)	2 (7.2%)
Written Texts 13–14	59	33 (56%)	9 (15%)	17 (29%)
Written Texts 16–17	67	33 (49.3%)	11 (16.4%)	23 (34.3%)
Written Texts 19–21	39	20 (51.3%)	6 (15.4%)	13 (33.3%)
Written Texts Adults	118	52 (44.1%)	22 (18.6%)	44 (37.3%)
Parental Speech to Toddlers 1;8–2;2	146	65 (44.5%)	38 (26%)	43 (29.5%)
Children's Storybooks	219	94 (43%)	39 (17.8%)	86 (39.2%)
Entire Database	384	153 (39.8%)	82 (21.4%)	149 (38.8%)

and the newer system of *Pi'el-Pu'al-Hitpa'el*, with both sub-systems expressing the full range of *binyan* semantic-syntactic functions (Ravid 2019; Ravid et al. 2016; Schwarzwald 2002). Given the research background described above, *family composition* in terms of the two sub-systems is a measure of morphological and lexical distance, i.e., morpho-lexical diversity. Morphological and lexical distance is lower among members of the same sub-system for two reasons. First, as each of the two sub-systems shares specific morpho-phonological characteristics (see Sect. 1.4.1); and second, as a root-sharing family within the same sub-system often tends to be more lexically uniform than across the two systems, e.g., *nixnas* 'enter', *hixnis* 'make-enter, insert', and *huxnas* 'be inserted' (older sub-system), versus *kines* 'gather,Tr.', *kunas* 'be gathered', and *hitkanes* 'gather,Int'. Therefore, the composition of a root-sharing derivational family across the two sub-systems indicates a greater morpho-lexical diversity. Moreover, the very number of families composed of the newer (*Pi'el*-based) sub-system is also an indication of morepho-lexical diversity, as this sub-system is the habitat of lexical productivity and innovation (Bolozky 2009; Laks 2013; Ravid 2019).

Table 15 presents the composition of families across the study's age groups and in the entire database by number and percentages of verbs in the first, the second, or both sub-systems. For example, toddler speech has 31 different families, 28 of which (90%) consist of conjugations within a single sub-system, and 3/4 of which are *binyan* conjugations in the older sub-system of *Qal-Nif'al-Hif'il-Huf'al*.

What is shared across virtually all age groups in Table 15 is (i) the fact that the overwhelming majority of families are contained within one sub-system; and (ii) the dominance of the older sub-system—accounted for by the great prevalence of *Qal* in Hebrew in general, and language development in particular—with its associated *binyan* conjugations. Like the frequency of *Qal*, the dominance of the older sub-system relatively declines with age and literacy, but is always greater than either of the two other options. This is one more piece of evidence that *Qal* serves as the launching board for the emergence of verb learning in general and verb derivational verb families as the focus of the current section. Only in the compiled database are the three options more equally distributed (almost 40% of the older sub-system and of integrated families, and slightly over 20% of the new sub-system).

Beyond these shared distributions, Table 15 shows three roughly distinct patterns—early spoken child language, interim spoken and written child language, and adolescent and adult production. In the youngest child speech corpora, up to age 4, families are few, and they are composed of solely the older sub-system (80–90%), while families composed of solely the newer sub-system are few (under 10%), and there are almost no families composed of the two sub-systems. The overwhelming composition of most families is either *Qal* / *Hif'il* (*nafal* / *hipil* 'fall / drop,Tr') or *Qal* / *Nif'al* (*gamar* / *nigmar* 'finish / be all gone'). In corpora of children up to age 10, in contrast, the number of families in the older sub-system declines to about 70%, while those of the newer sub-system and/or comprising both sub-systems rise. A frequent alternation of *Pi'el* / *Hitpa'el* is added in these years (*sovev* / *histovev* 'turn,Tr / turn around'). In the spoken and written corpora of adults and adolescents, and especially in written texts, the older sub-system occupies 40–50% of the families, while the rest is shared by the newer sub-system (e.g., *bikesh* / *hitbakesh* 'ask / be asked') and by families composed of both (*hiksha* 'make hard' in *Hif'il* / *hitkasha* 'become hard / struggle' in *Hitpa'el*). This change takes place side by side with the exponential growth in number of derivational families and the increase in family size, together with the introduction of new, lexically specific items, often in families sharing an abstract root.

By the time derivational families spread over both sub-systems in linguistically mature users, connections are forged not only between the frequent pairs of *Qal* / *Nif'al, Qal* / *Hif'il* and *Pi'el* / *Hitpa'el*, but also between active and passive conjugations within the same sub-system (e.g., the *Hif'il* / *Huf'al* pair *hidgish* / *hudgash* 'emphasize / be emphasized'), and between members of both sub-systems, as in the example above. Later patterns of frequency start emerging, as between basic *Qal* and middle voice *Hitpa'el* (*yashav* / *hityashev* 'sit down / seat oneself' in written 9–10 texts); or causative *Hif'il* and *Hitpa'el* (*herim* / *hitromem* 'lift / rise up' in the storybooks corpus). Moreover, the semantic relationship between family members within and across the sub-systems becomes less predictable and more diverse with age and schooling. While semantic coherence is the topic of the next section, the increasing diversification of semantic relations between the same *binyan* conjugations is worth noting here. For example, the frequent *Qal* / *Nif'al* link is restricted to a basic / middle voice relationship, as in *shavar* / *nishbar* 'break,Tr / break, Int' in the peer talk of 2;3–3 year olds; but the written texts of 13–14 year olds reveal the same *Qal* / *Nif'al* link, this time expressing a basic / passive voice relationship in *garam* / *nigram* 'cause/be caused'. Older age groups exhibit usage of less frequent relationships across the sub-systems, such as basic, juvenile *ne'evad* 'go missing' in *Nif'al* together with *Pi'el* *ibed* 'lose' in the 7–8 peer talk.

6.4.1 Interim conclusion

The changing patterns of family distributions across the sub-systems support the emergence of Hebrew verbs as both lexical and morphological entities. To begin with, a limited number of small-sized families with low morphological distance among their members, restricted within a specific sub-system, occur in younger speech corpora, sustaining the learning of the core verb lexicon and highlighting core transitivity relations among these verbs. With age and schooling, families not only grow bigger and more numerous, but their members also become more morphologically distant, reflecting a subtler and richer lexical and *Aktionsart* composition by incorporating members across the two systems. The older system continues to grow in terms of rarer, higher register families and more lexically specific members (*darax / hidrix* 'step / guide' in *Qal* and *Hif'il* respectively, written texts by young adults). At the same time, members from the newer system introduce new semantic permutations of the basic semantics (*ne'elam / he'elim / hit'alem*, get lost / hide, Tr / ignore' in written texts by 11th graders). It is only in the adults' corpora—spoken, written and expert—that the often-cited lexically creative and diverse Hebrew derivational families combining both sub-systems start to emerge. This is illustrated by the nice example of the *q-d-m*-based family represented by *kadam / hikdim / kidem / hitkadem* 'precede / make earlier / promote / make headway' (*Qal, Hif'il, Pi'el* and *Hitpa'el* respectively) in the adults' written texts.

6.5 Family coherence

While we have been treating *binyan* conjugations from both structural and semantic aspects, our discussion of roots in verbs and root-related families has related so far to the root as a primarily structural entity. In the introduction to Part I above (including illustrations in Tables 4 and 5) we showed why defining a derivational family based on semantic relatedness between root-related verb members would soon drive this analysis into a hopeless quagmire (Ravid et al. 2016). However, it cannot be denied that semantics plays an important role in Hebrew speakers' conceptualization of root relations (Berman 2012; Frost et al. 1997; Ravid 2003; Schwarzwald 2001). What follows below is a first attempt at examining the degree of semantic coherence in the root-related families. Recall that 60% of the roots in our entire database were singletons, i.e., one root=one verb (Table 12). Singleton verbs uphold semantic coherence in the sense of presenting Hebrew users with a constant and consistent meaning associated with the same root across different temporal categories. It is the remainder, family-relating component of the root inventory in our study that required this semantic investigation.

Thus, the purpose of the final analysis regarding *family* coherence was to determine the extent to which members in the database verb families were semantically related. To this end, the Methods section above (Part I, Sect. 3.2.5) describes the process whereby a list of 707 root-sharing verb pairs was created in the entire corpus, where each verb was paired with all other verbs sharing the same root skeleton. This list was presented to 64 native-speaking experts in Hebrew developmental psycholin-

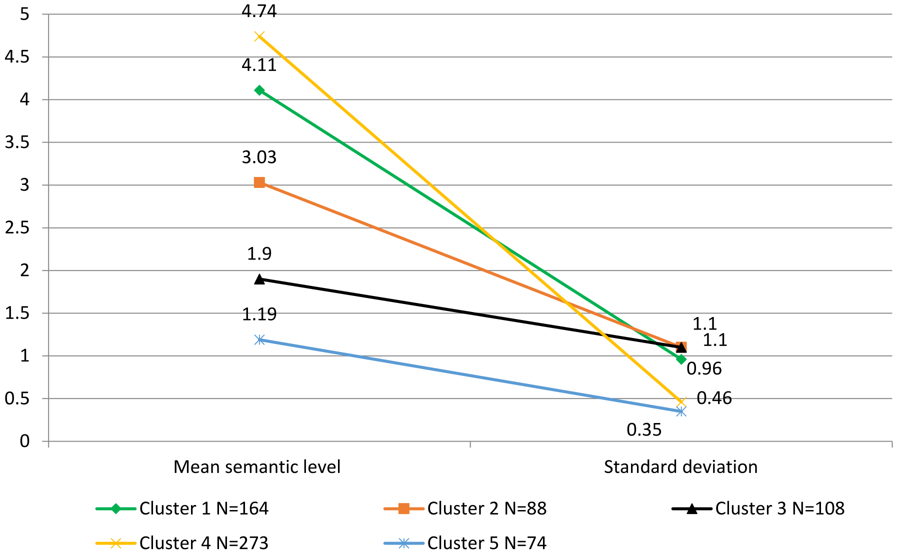


Fig. 4 Five latent clusters depicting semantic level and degree of agreement among judges resulting from the LCA procedure (Color figure online)

guistics, who ranked each pair on a scale of 1–5, with 1 indicating no meaningful relationship between members of the pair, and 5 indicating a strong semantic relationship. The pairs were grouped into five clusters or levels of semantic relatedness by a Model Based Latent Class Analysis (LCA) procedure. A Model Based Latent Class Analysis (LCA) enables the identification of unobservable subgroups that are similar, based on observed characteristics—in the current case, mean semantic coherence ranks for each of the verb pairs.

Figure 4 presents the results of the cluster analysis on the mean semantic level and standard deviations (SDs) of semantic relatedness and agreement among judges of root pairs in the entire database. It shows that most root pairs in the database (62%, including Cluster 4, 273 pairs and Cluster 1, 164 pairs respectively) were semantically coherent—above level 4. Next was Cluster 2 with 88 pairs (12%), expressing a middling level (3) of semantic coherence. Finally, 26% of the root pairs (Cluster 3, with 108 pairs, and Cluster 5, with 74 pairs respectively) expressed a lower level (under 2) of semantic relatedness. These results are strongly related to the fact that most families in the database consisted of two members, in most cases from the same sub-system. These conclusions are supported by the analysis depicted in Table 16 and Fig. 5, presenting a simple linear regression analysis to test if the family size significantly predicted mean semantic relations ($R2 = 0.11$, $F(1, 691) = 81.32$, $p < 0.001$), such that large families were expected to have lower mean semantic relations ($\beta = -0.44$, $p < 0.001$). This analysis indeed indicated that the larger the family, the lower the degree of semantic relatedness among its members. Below we analyze the semantic coherence of verbs in small, two-member families (Sect. 6.5.1), and in larger families (Sect. 6.5.2).

Table 16 A simple linear regression analysis of family size as predicting semantic coherence

	Model 1
(Intercept)	3.72***
	−0.04
Family size	−0.36***
	−0.04
N	693
R^2	0.11

*** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$.

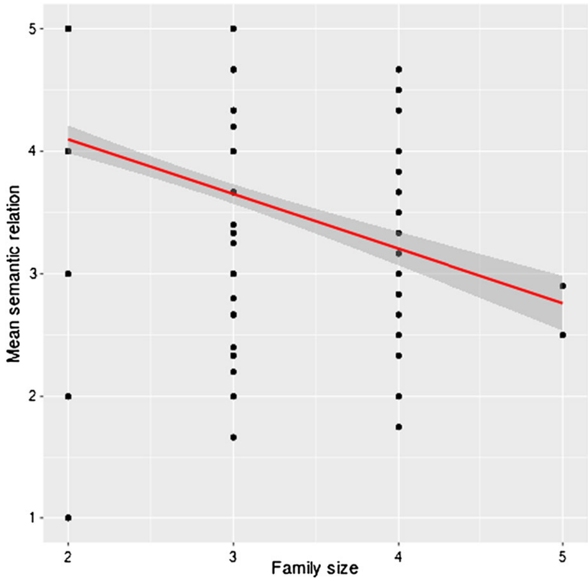


Fig. 5 The relationship between family size and family coherence (simple linear regression analysis)

6.5.1 Families with two members

We first present results of the analysis of semantic coherence in two-*binyan* families, which constituted the overwhelming majority of the families in the entire database.

Table 17 presents degrees of semantic coherence in two-member families in the entire database as well as in each age-related corpus, showing the number of families that were assigned each coherence level and the mean coherence level (score) per corpus. Importantly, the bottom line of Table 17 shows the mean coherence levels assigned to families whose members may have occurred separately in different corpora within the entire database. An example of such a pair is the *ʔ-z-n* family (cf. *ózen* ‘ear’, *moznáyim* ‘scales’) with *he’ezin* (*Hif’il*) ‘listen’, which first occurred in the children’s storybooks, and (*Pi’el*) *izen* ‘balance’, which first occurred in the written (13–14) texts.

Coherence levels across the entire database revolved around Level 4, as 80% of the two-*binyan* families in the database were assigned the highest levels of 4 and 5—e.g., *cilem* / *hictalem* ‘take photo / get photographed’ (5), or *yixes* / *hityaxes* ‘attribute / treat’ (4). Our first conclusion is that two-*binyan* families in the database were highly

Table 17 Degrees of semantic coherence among pairs of root-related verbs in two-*binyan* families across the database. The first column lists the total number of two-*binyan* families in the relevant age group. The next five column lists the number of two-*binyan* families per each level of semantic coherence, from 5 (high semantic coherence) to 1 (no semantic coherence)

Age Group	# of two-*binyan* families	Level 5	Level 4	Level 3	Level 2	Level 1	Average Score
Toddler Speech 1;8–2;2	28	20	2	1	2	3	4.2
Peer Talk 2–2;6	4	3	–	1	–	–	4.5
Peer Talk 2;6–3	10	6	2	2	–	–	4.4
Peer Talk 3–4	16	11	2	2	–	1	4.37
Peer Talk 4–5	19	11	5	1	–	2	4.21
Peer Talk 5–6	24	14	4	4	1	1	4.21
Peer Talk 7–8	20	8	6	4	–	2	3.95
Written Texts 9–10	26	21	3	1	1	–	4.7
Written Texts 13–14	51	26	14	2	4	5	4.02
Written Texts 16–17	56	26	17	4	4	5	3.98
Written Texts 19–21	33	17	6	5	4	1	4.03
Written Texts Adults	94	43	22	11	11	7	3.88
Parental Speech to Toddlers 1;8–2;2	130	83	19	10	8	11	4.16
Children's Storybooks	156	92	38	9	17	19	4.06
Entire Database	273	147	65	22	25	14	4.12

coherent. However, as root-related pairs are derivational entities, the general average score did not exceed 4 by much, reflecting the typical degree of unpredictability associated with derivation (Ravid 2019). For example, the *he'ezin* / *izen* pair above was rated with the semantic coherence level of 1, i.e., a virtual absence of semantic relatedness. This reflects the fact that the high-register *Hif'il* verb 'listen' comes from Biblical Hebrew, whereas the *Pi'el* verb *izen* 'balance' is a new derivation, expressing a scientific concept. That is, semantic opacity may be associated with a new verb being coined in the *Pi'el*-based sub-system from an old root prevalent in the *Qal*-based sub-system.

Development. From a developmental point of view, examining the changing patterns within and across separate corpora, coherence levels were higher in younger age groups—e.g., *paxad* / *hifxid* 'fear / frighten' (5), 2;6–3 peer talk; or *zaz* / *heziz* 'move,Int / move,Tr' (5), 5–6 peer talk. Not only did levels concentrate on the higher (left) side of Table 17 in younger age groups up to age 9–10—most of these corpora showed missing lower values, most probably due to the small number of families they each contained. For example, the peer talk group of 2;6–3 year olds had 10 2-member families altogether, of which 8 were rated as levels 5 and 4, and two families were assigned level 3. No families were assigned levels 1 or 2 in this age group.

With age and schooling, this picture becomes more balanced. As the number of two-member families increased in older corpora, pairs were assigned all levels of coherence—although 4 and 5 were still the most frequent across the board. Mean degrees of coherence slightly declined in the older age groups, side by side with the increase in family frequency, and especially in written texts, where all levels of coherence were represented. This reflects the increased number of verbs sharing the same root skeleton with widely diverging semantics, e.g., *ana* / *ina* 'respond / torture' (1, sharing skeleton *'-n-y* in the written texts of adults), or *henec* / *nacac* '(sun) rise / shine' (2, sharing skeleton *n-c-c* in adults' CDS to toddlers).

Most pairs with very low semantic scores (levels 1 and 2) were found in the older age groups. For example, the two members of the pair *lakax* / *hitlaké'ax* 'take / catch fire' (level 1) first occurred together in the written narratives of adults. Such pairs reflect the long history of Hebrew, where phonological changes and semantic shifts have resulted in same-skeleton verbs sharing little semantics. This is illustrated by the following pairs occurring in the children's storybooks: *hirgish* / *hitragesh* 'feel / grow excited' (level 3) (root *r-g-š*), *hexezik* / *xizek* 'hold / strengthen' (2) (root *ħ-z-q*), *bara* 'create' / *hivri* 'get well' (2) (root *b-r-ʔ*), and *cava* 'paint' / *hicbi'a* 'point' (1) (root *c-b-'*). Finally, highly opaque defective roots were assigned lower semantic coherence levels, e.g., *naga* / *higí'a* 'touch / reach' (2) based on root *n-g-'* (toddlers' CS), or *hoxí'ax* / *hitvaké'ax* 'prove / argue' (2) in the young adults' written texts, based on root *y-k-ħ*. A similar picture of decreasing transparency with age and schooling has recently been shown for German and Italian diminutives (Dressler et al. 2019).

6.5.2 Semantic coherence in larger families

While two-member families constituted the majority of derivational verb families in the entire database, it also contained 111 families with more members (11% of all roots). This made the assessment of semantic coherence a complicated task, as coherence was measured within each pair of verbs sharing the same root, family size notwithstanding. As explained above, all possible pairs were lined up, and each pair was given a separate score. For example, root *k-n-s* had a family of three members in the adults' written texts—*nixnas* 'enter', *hixnis* 'introduce, make come in', and *kines* 'gather,Tr'. The pair *nixnas* / *hixnis* was assigned level 5, *hixnis* / *kines* was rated as 3, and *nixnas* / *kines* was assigned level 2. Our analyses here relate to the entire database, as Table 12 makes it clear that larger families of three and four members hardly occurred even in the older age groups and written language.

To assess semantic coherence in larger families of three, four, and five members, containing numerous pairings of verbs, a compiled measure of semantic *uniformity* was developed across different pairs within a family. To determine to what extent the semantic rating of the pairs within a single root-based family was *uniform*, the distance between the semantic ratings of these pairs was calculated. A family was considered to be semantically uniform if the distance between pair ratings covered a small segment of the semantic scale, e.g., 4–5 or 3–4.

Four uniformity groupings were thus identified in the entire large-family database: (i) *uniform* families with highly coherent semantics, i.e., where all pairs were rated 4–5—e.g., *he'ir* / *orer* / *hit'orer* 'wake up,Tr / arouse / wake up,Int'); (ii) *uniform* families with middle coherence levels, where most pairs were rated 3–4—e.g., *asak* / *he'esik* / *hit'asek* 'be occupied with / employ ∼ occupy / be involved ∼ fiddle with; (iii) *non-uniform* families where pairs were semantically rated from 2–5, e.g., *hiskim* / *sikem* / *histakem* 'consent / add up,Tr (also 'summarize') / add up to'); and (iv)

Fig. 6 Distribution of the four uniformity categories in the 89 three-member families across the entire database (Color figure online)

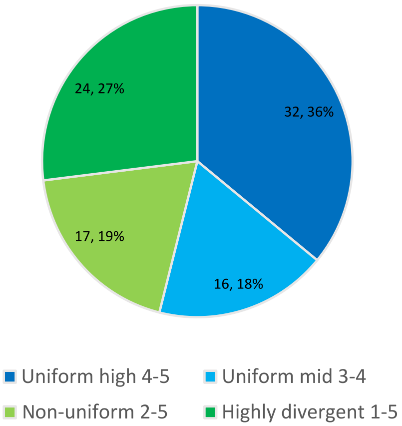
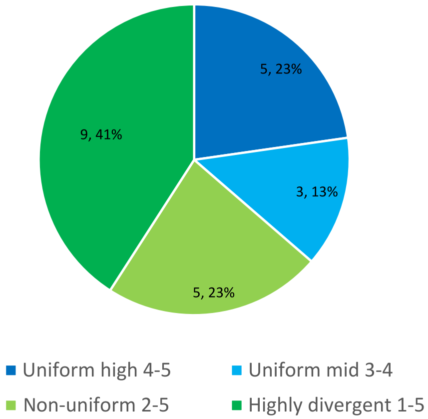


Fig. 7 Distribution of the uniformity categories in the 22 largest families (four and five members) across the entire database (Color figure online)



highly divergent families where pairs were rated across the full semantic coherence scale from 1–5, e.g., *pana* / *hifna* / *hufna* / *pina* / *hitpana* ‘turn to / refer / be referred / evacuate,Tr / evacuate,Int (also turn one’s mind to)’. Note that the potential category of uniform 1–2 ratings was not identified in the large families. That is, no family with three or more members was found such that all pairs were rated as non-semantically coherent. Recall that this was not true with regards to smaller, two-member families, where 39 out of the 273 two-member families (14%) were rated 1–2 (e.g., *rigel* ‘spy’ / *hitragel* ‘become habituated’).

Figure 6 depicts the distribution of the four uniformity categories in the 89 three-member families across the entire database. It shows that roughly half of these families were uniform (categories i and ii), and roughly half were non-uniform (categories iii and iv). Figure 7 depicts the distribution of the uniformity categories in the 22 largest families consisting of four and five members. It shows that 2/3 of these families were non-uniform with categories iii and iv (i.e. divergent), while only 1/3 of them were uniform, with categories i and ii. The family based on *p-t-ħ*, which had four members in the entire database, illustrates non-uniform relations. The six pairs deriving from this family were rated as follows: *patax* / *niftax* ‘open,Tr / open,Int’—5;

patax / *pité'ax* 'open,Tr / develop,Tr'—2; *patax* / *hitpaté'ax* 'open,Tr / develop,Int'—2; *niftax* 'open,Tr / *pité'ax* 'open,Int / develop,Tr'—2; *niftax* / *hitpaté'ax* 'open,Int / develop,Int'—2; *pité'ax* / *hitpaté'ax* 'develop,Tr / develop,Int'—4. Figures 6 and 7 thus indicate that the larger the root-based family, the less semantically uniform it tends to be, as larger families in our databased tended to contain members that diverged in their semantics.

Taken together, the analyses of semantic coherence and uniformity in Sects. 6.5.1 and 6.5.2 underscore the role of the root in the verb lexicons of older, more literate, Hebrew users. Mature Hebrew discourse has a larger proportion of root-based families in general, and a larger proportion of large families in particular, indicating a tight organization by roots as morphological entities. But, at the same time, mature verb lexicons are less dependent on the root as a semantic entity, as evidenced by the fact that larger root-based families are largely non-uniform—that is, they display more semantic diversity and lexical specificity than smaller families. The 14% of two-member families whose members hardly share semantic ties enhance this impression. However, even in mature, literate speaker / writers' lexicons root organization is not merely structural, as evidenced by the absence of 'non coherent uniformity' (three or more pairs within the family all rated 1 or 2) in larger families.

7 Concluding general discussion

This paper presents a psycholinguistic analysis of the acquisition and development of Hebrew verbs from toddlerhood to adulthood, based on a database of about half a million words compiled from spoken and written corpora produced by native Hebrew users. The study is grounded in the Semitic Hebrew typology, whose prototypical expression is morphology, and specifically, the Hebrew verb system. Across the database, and within all the corpora that make it up (in the development / literacy sequence first presented in Table 7), all derivational components of the Hebrew verb system were analyzed quantitatively and qualitatively in types and tokens. The analyses in Part I presented the distributions and characteristics of verb lemmas, roots (including structural root categories), and *binyan* conjugations (including the two subsystems). Part II focused on root-based verb derivational families in terms of family frequency, family size, family composition and the semantic coherence of families. Grounded in the Usage-Based Approach to psycholinguistics, this study was able to provide new empirical evidence from natural discourse regarding the emergence and consolidation of the Hebrew verb system. Novel structural and semantic analyses of the data have made it possible to account for the role of the Semitic notions of *root* and *binyan* and the networks they create in Hebrew verb learning.

According to the usage-based account of language learning, grammatical systematicity emerges from language use, so that older and more experienced speaker / writers are able to construe items within the systems that they make up (Ackerman et al. 2009; Diessel and Hilpert 2016). The previous Sects. 1–6 presented the current state of the art regarding the acquisition and consolidation of different facets of the verb system along the developmental axis. This concluding section takes two over-arching views of the Hebrew verb system. First, from the bird's eye vantage point of the verb

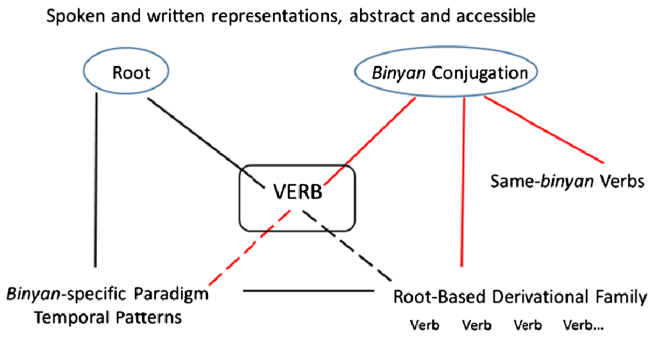


Fig. 8 Hebrew verbs in their morphological networks: A bird's eye view

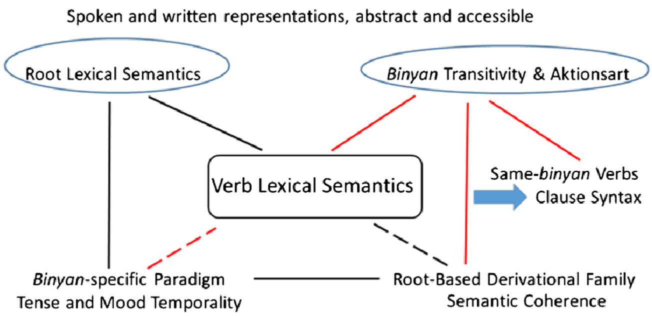


Fig. 9 The lexical semantics network of Hebrew verbs: A bird's eye view

in the mature system; and second, a developmental review taking into account how Hebrew verbs evolve from the confluence of factors described in the current paper.

7.1 'Lexical quality' in the verb system

Based on the literature reviewed in the introduction and the entire set of analyses described in the two parts of this paper, Figs. 8 and 9 represent the kind of knowledge that has been elsewhere termed 'lexical quality' (Perfetti 2007).[24] In the current context, this term is taken to refer not only to command of the form and full lexical semantics (including ambiguity, homonymy and homophony) of each verb; but also to the systematic command of the formal organization it is couched in, and the different types and degrees of lexical and categorial semantic properties typical of this organization. 'Lexically qualitative' knowledge of the verb and the verb system is, on the one hand, fully automatic in adult native speakers, allowing easy manipulation and retrieval of all related words and morphemes; but at the same time it is also highly abstract, involving mostly written, metalinguistic representations (Ravid and Schiff 2006b, 2012) of the type termed E2/E3 by Karmiloff-Smith (1992). It is this dense

[24]The notion of 'lexical quality' (Perfetti 2007) initially comes from the reading literature. It refers to the idea that words lie at the heart of reading comprehension, so that knowledge of the full range of meanings, structures and usages of each word is critical to understanding a text containing that word.

and complex, deeply embedded yet accessible knowledge of lexical morphology that underlies Hebrew adolescent and adult creativity in coining and comprehending new words, and especially in the fluid, smooth manipulation of root skeletons across lexical items (Ravid 2003).

Figures 8 and 9 depict the inherent role of morphology in the mature configuration that we call 'the Hebrew verb system' from structural and semantic perspectives, respectively. Both figures show that beside being a lexical item with its own unique content and contexts, each Hebrew verb is embedded in a network of which it forms an integral part (Levie 2012). Take for example the two verbs *ganav* 'steal' and *hitganev* 'move stealthily, sneak in', two autonomous verbs with their own lexical semantics. Figure 8 provides the morphological context for their association. These two verbs share the same root skeleton, *g-n-b*, in two different *binyan* conjugations—*Qal* and *Hitpa'el* respectively (top of Fig. 8), and thus form part of a verb derivation family (bottom right), together with other verbs such as *nignav* 'be stolen', *higniv* 'bring in stealthily', and *hugnav* 'be brought in stealthily'. By virtue of belonging to their respective *binyan* conjugations, each of them shares structure with other *Qal* and *Hitpa'el* verbs (on the right of Fig. 8). Finally, each verb lemma represented here by *ganav* and *hitganev* (respectively) in fact constitutes a set of temporal stems sharing the same root with *binyan*-specific patterns (bottom left)—e.g., *ganav* 'stole', *gonev* 'steals', *yignov* 'will steal', *gnov* 'steal,IMP', and *li-gnov* 'to-steal' for 'steal'. The 'lexically qualitative' knowledge of verbs represented in Fig. 8 includes not only the fully automatic morphological system and the specific morpho-phonological permutations of verb forms, but also, and especially in adults, the written forms of these verbs and their roots—including the vowel writing norms that support the abstract representation of a root such as *g-n-b* in its written Hebrew form גנב (Ravid 2012).

Figure 9 represents the verb and its morphological environment in semantic terms, based on the same examples. The two verbs share a notion of stealing (perceived as root lexical semantics, top left), however *ganav* in *Qal* stands for the basic notion, while *hitganev* in *Hitpa'el* expresses more complex lexical semantics (middle of chart). Importantly, Fig. 9 shows (top, right, and middle of chart, right) that *binyan* shift to *Hitpa'el* involves not only a change in *Aktionsart* (durative temporality) but also of transitivity values—in this case, from transitive 'steal' to intransitive, ergative 'steal in', thereby a reduction in number of arguments. This part of the chart highlights what it means to command the system of *binyan* conjugations in terms of its role in the construction of clause syntax. The *binyan*-specific paradigm of each verb (bottom, left) provides multiple opportunities for Hebrew users to experience/produce the root skeleton in different temporal patterns but with the same lexical substance. Other verbs in the derivational family share the notion of stealing, with different degrees of semantic coherence or transparency (bottom, right). In this specific context, lexically qualitative knowledge of the *g-n-b* family also includes the associated slang notion of 'being cool'.

One direct outcome of this morpho-lexical organization is the well-known ability of Hebrew speaker/writers to extract root skeletons and re-insert them efficiently in new content words (Ashkenazi et al. 2019; Berman 2000, 2003, 2012; Bolozky 1999; Frost et al. 1997, 2000; Laks 2013; Levie et al. 2017, 2019; Ravid 1990, 1995, 2003; Ravid and Bar On 2005; Ravid et al. 2016; Schwarzwald 1981, 2000, 2001). In consequence of being organized in root-related families, Hebrew words readily lend themselves to the extraction of root skeletons, which, in turn, serve in new words. Root extraction is an extremely accommodating process in Hebrew which is fed by any and all word categories—content or grammatical, Hebrew or foreign, with or without internal morphological structure, mostly targeting verbs as the innovative lexical items (Bolozky 2004; Laks 2018; Nir 1993; Ornan 2014; Ravid 2019; Schwarzwald 1981). The process identifies and extracts the consonantal skeleton of the base word, combining it with the appropriate category-assigning pattern, resulting in a word that conveys the meaning of the base word (or a facet thereof), much like zero derivation in English (Clark and Clark 1979). The non-developmental Hebrew literature—linguistic, psycholinguistic and typological—teems with examples of new-verb derivation such as *tizmer* 'orchestrate' from *tizmóret* 'orchestra' (Bolozky 2003; Ephratt 1997; Kassovsky 1985; Ravid 1990; Schwarzwald 2000, 2001; Schwarzwald and Neradim 1995).

Against this discussion of roots in the literature and in our study, recall the stem/word-based approach that denies the status of the root as a morpheme, ascribing this status to the *binyan* template (Bat-El 1994, 2003, 2017; Ussishkin 2005). Kastner (2019), while working within the same theoretical framework as the above-mentioned authors, reaches the inverse conclusion: He argues that it is the root that is an independent morpheme while *binyan* templates are epiphenomenal. Kastner presents a number of arguments against the stem-based theory of Semitic morphology. One of them is that this theory predicts that there must always be a *CaCaC* form to use as a base, which is simply not true, given that the verb system is derivational with inherent gaps. Thus, according to Kastner, having one root as the base of derivation for all forms is a more useful generalization. Moreover, stem-based analyses are limited to third person singular past tense forms (the citation form), ignoring the lack of psycholinguistic evidence to support this assumption. Finally, Kastner claims that the stem-based approach attributes an exaggerated role to templates, ignoring their syntax and semantics—as shown by our findings in the current paper.

Returning to our bird's eye depiction of the facile, creative, automatic derivation processes involving roots and patterns, note that they are restricted to mature, literate knowledge. Specifically, virtually all of the published literature on Hebrew-speakers' ability to manipulate roots, *binyan* conjugations and derivational families[25]—experimental, discursive, or Internet-based—involved educated adults from mid-high socio-economic status (SES). The overview offered by Figs. 8 and 9 is no exception. The explicit goal of this paper was to detect the developmental process that leads up to this mature, literate system, as presented in Fig. 10.

7.2 The route to the derivational verb family in Hebrew

Little of the morpho-lexical knowledge of verbs, their morphological components and derivational families depicted in Figs. 8 and 9 is at the disposal of young Hebrew-speaking children. The construal of the verb system even by the end of elementary school is very different from the abstract, complex, dense and accessible multi-modal properties of the mature system. Based on the empirical evidence presented in the

[25] Excepting Levie et al. 2017; Ravid 1995 and Schwarzwald 1981.

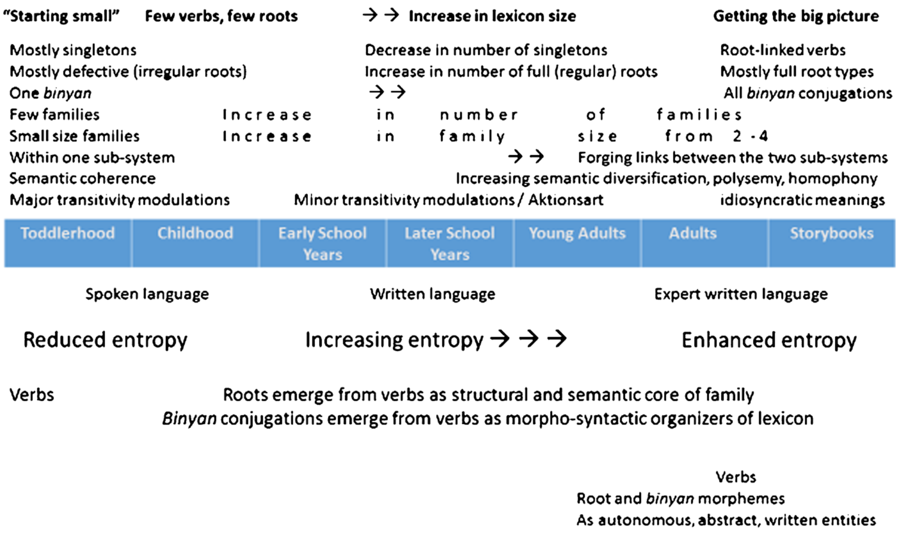


Fig. 10 Emergence and consolidation of the Hebrew derivational verb family

current article, our main claim is that roots, *binyan* patterns and derivational verb families are all *emergent properties* of the verb system as it develops in variegated communicative contexts.

Figure 10 presents the bird's eye view of the emergence and consolidation of the verb system, using the notion of *entropy*—degree of unpredictability or average information content (Gray 2011). From left to right, it shows how the patterns we detected in the data shift along the developmental and modality axis (depicted at the bottom of the chart) from low or reduced entropy in the spoken discourse of toddlers and preschool children to increasing entropy in school-going populations, alongside the increased reliance on written language; and enhanced entropy in the speech and writing of adolescents and adults. This summarizing chart depicts the converging patterns that together enable the informational content of verbs to increase and diversify along these axes, with root and pattern systematicity as the by-product of this entropy increase.

Starting from the left side, reduced entropy means that even in Hebrew, a Semitic language, the morphological verb system "starts small" (Elman 1993). Initially, verb learning is indeed verb learning: Most contribution to growth in the verb lexicon is made by singleton verbs—overwhelmingly, single verbs based on single roots. In early toddlerhood and childhood, there appears a strongly lexical verb core of mostly singleton items in one *binyan*—the *Qal* conjugation. Morphology is embarked upon by usage of a restricted, then growing number of temporal (and agreement) verbforms within *Qal* (Ashkenazi 2015). Initial morphological manipulations involve teasing apart the root and pattern components from highly opaque verb forms (Lustigman 2016) based on defective (irregular) roots, facilitated by highly transparent single-lemma lexical semantics (Ashkenazi et al. 2016, 2019; Ravid et al. 2016). Morphological learning further rides on a small number of two-*binyan* families, virtually all set within a single sub-system—mostly the older, *Qal*-based sub-system—and subsequently with highly coherent semantics. These few and small families introduce the major inchoative and causative transitivity modulations, ushering in basic clause syntax (Berman 1985). Learning verb semantics helps children construe transitivity values, attributing them to *binyan* pairs prevalent at this time, prior to the consolidation of derivational families.

The middle part of the chart focuses on the elementary school years, showing how the notions of root and *binyan* emerge from repeated structural similarities and semantic modulations in the rapidly increasing verb lexicon of child Hebrew—side by side with opening literacy channels by reading and writing instruction in the school system. The growing number of verbs based on full (regular) radicals—many of them from written language—makes it possible to discern consonantal root skeletons from vocalic patterns in speech and writing and to detect similarities among an increasing number of verbs sharing roots and patterns. Construing non-linear structure is moreover enabled by the increase in verbs based on different *binyan* conjugations in addition to *Qal*. The notion of the *derivational verb family* is a further emergent property of the evolving system, essentially dependent on the rise in number of pairs of verbs sharing a discernible root, and the initially restricted appearance of families with more than two members. This is the watershed period, when the overwhelmingly singleton lexicon first acquires its derivational organization (Ben Zvi and Levie 2016; Ravid 2003).

This is also the time when functions typically associated with the *binyan* conjugation system productively emerge. Consider, for example, the transitive-causative relationship, which is highly prominent in early childhood in a small number of verb pairs (Sect. 5.4 above). A qualitative analysis of verbs occurring in our database corpora indicates that productivity of this relationship is manifested only from age 5 onwards in the appearance of a large number of new, highly transitive *Hif'il* verbs. These are concrete verbs like *he'exil* 'feed', *heki* 'vomit', *he'emid* 'make stand', *he'ela* 'make go up', *hexbi* 'hide', *he'evir* 'move, Tr', *hifsik* 'stop', *hirbic* 'hit', *hishpric* 'spurt', *he'ir* 'wake,Tr', *higdil* 'make big', and *he'if* 'make fly'; and many verbs of social interaction, including mental and saying verbs, e.g., *hikir* 'recognize', *hecik* 'annoy', *hifri'a* 'cause interference', *hirgi'a* 'calm,Tr', *hicxik* 'make laugh', *hirsha* 'allow', *hig'il* 'cause disgust', *he'enish* 'punish', *he'eliv* 'offend', *he'eshim* 'accuse', *hishpil* 'humiliate', *hidrix* 'instruct', *hishpi'a* 'influence', *higish* 'present', *hodi'a* 'announce', *hici'a* 'suggest', *hisbir* 'explain', *he'etik* 'copy', and *hifna* 'refer to,Tr'. Abstract and lexically specific high-register transitive and causative *Hif'il* verbs appear in yet older age groups, e.g., *higdir* 'define', *hevix* 'embarrass', *hegiv* 'react, *hidgish* 'emphasize', *hikpid* 'make sure', *hit'im* 'make fit', *hishlim* 'make whole', *hishlix* 'jettison', *hinmix* 'make low', *hosif* 'add', *hexdir* 'insert, *hit'im* 'make fit', *hishki'a* 'invest', *hefic* 'disseminate', *hovil* 'lead', and *hesit* 'incite'. Latest appearing *Hif'il* verbs tend to be less agentive, e.g., *hivxin* 'notice', including, later on, a small group of inchoative *Hif'il* verbs such as *hismik* 'blush' or the ambiguous *hirxik* 'go far / make far'. Taken together, the syntactico-semantic, categorial functions of the *binyan* consolidate and diversify at a time of great lexical learning in elementary-school children (Anglin 1993).

Figure 10 shows that it is only by late adolescence and adulthood that Hebrew speaker / writers are endowed with the celebrated, morphologically and lexically diverse Semitic lexicon that enables so much new-verb creativity. This mature lexicon with enhanced entropy (i.e., a great load of information) has multiple (and larger) derivational verb families (though fewer than previously thought); and much of it is based on regular, transparent roots, often derived from words in other lexical classes. Larger derivational families encompass the two subsystems and therefore have diverse semantic and structural properties, including ambiguities of many kinds—synonymy, homonymy, and also homography (Bar-On et al. 2017). The abstract notion of the Semitic root as a written entity linking many verbs in a derivational family, prevalent in the educational, linguistic and literacy literature, is the meta-linguistic outcome of this converging knowledge.

The current database, though unique in size and diversity, cannot be said to fully represent all Hebrew contexts. It has a spoken bias and lacks a proper representation of expert-written texts. Therefore, it might under-represent transparent roots and larger root-based families: we can assume that in the same way that half of the singletons in this study had 'hidden' family members outside our database, two-member (and larger) families in this corpus may are actually be larger. From the point of view of the naïve Hebrew learner, each singleton and each family constitute the portal to more, root-related verbs, nouns, and adjectives.

Acknowledgement This study was supported by Israel Science Foundation grants 285/13 and 219/17 to Dorit Ravid.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

Abbot-Smith, K., & Tomasello, M. (2006). Exemplar-learning and schematization in a usage-based account of syntactic acquisition. *The Linguistic Review*, *23*, 275–290.

Ackerman, F., Blevins, J. P., & Malouf, R. (2009). Parts and wholes: Implicative patterns in inflectional paradigms. In J. P. Blevins & J. Blevins (Eds.), *Analogy in grammar* (pp. 54–82). Oxford: Oxford University Press.

Ackerman, F., & Malouf, R. (2013). Morphological organization: The low conditional entropy conjecture. *Language*, *89*, 429–464.

Aguado-Orea, J., & Pine, J. M. (2015). Comparing different models of the development of verb inflection in early child Spanish. *PLoS ONE*, *10*(3), e0119613. https://doi.org/10.1371/journal.pone.0119613.

Ambridge, B., Kidd, E., Rowland, C. F., & Theakston, A. L. (2015). The ubiquity of frequency effects in first language acquisition. *Journal of Child Language*, *42*(2), 239–273.

Anglin, J. M. (1993). *Vocabulary development: A morphological analysis*. *Monographs of the Society for Research in Child Development: Vol. 58* (p. 1). Chicago: University of Chicago Press.

Armon-Lotem, S., & Berman, R. A. (2003). The emergence of grammar: Early verbs and beyond. *Journal of Child Language*, *30*, 845–878.

Ashkenazi, O. (2015). *Input–output relations in the early acquisition of Hebrew verbs*. Unpublished doctoral dissertation, Tel Aviv University.

Ashkenazi, O., Gillis, S., & Ravid, D. (2019). Input–output relations in the early acquisition of Hebrew verbs. *Journal of Child Language* https://doi.org/10.1017/S0305000919000540.

Ashkenazi, O., Ravid, D., & Gillis, S. (2016). Breaking into the Hebrew verb system: A learning problem. *First Language*, *36*, 505–524.

Bar-On, A., Dattner, E., & Ravid, D. (2017). Context effects on heterophonic-homography resolution in learning to read Hebrew. *Reading & Writing*, *30*, 463–487.

Bar-On, A., & Ravid, D. (2011). Morphological decoding in Hebrew pseudowords: A developmental study. *Applied Psycholinguistics*, *32*, 553–581.

Bat-El, O. (1994). Stem modification and cluster transfer in Modern Hebrew. *Natural Language and Linguistic Theory*, *12*, 571–596.

Bat-El, O. (2003). The fate of the consonantal root and the binyan in Optimality Theory. *Recherches Linguistiques de Vincennes*, *32*, 31–60.

Bat-El, O. (2017). Word-based Items-and-processes (WoBIP): Evidence from Hebrew morphology. In C. Bowern, L. Horn, & R. Zanuttini (Eds.), *On looking into words (and beyond): Structures, relations, analyses* (pp. 115–134). Berlin: Language Science Press.

Behrens, H. (2006). The input–output relationship in first language acquisition. *Language and Cognitive Processes*, *21*, 2–24.

Ben Zvi, G., & Levie, R. (2016). Development of Hebrew derivational morphology from preschool to adolescence. In R. Berman (Ed.), *Acquisition and development of Hebrew: from infancy to adolescence* (pp. 135–173). Amsterdam: Benjamins.

Berman, R. A. (1978). *Modern Hebrew structure*. Tel Aviv: University Publishing Projects.

Berman, R. A. (1980). Child language as evidence for grammatical description: Preschoolers' construal of transitivity in Hebrew. *Linguistics*, *18*, 677–701.

Berman, R. A. (1981). Language development and language knowledge: Evidence from acquisition of Hebrew morphophonology. *Journal of Child Language*, *8*, 609–626.

Berman, R. A. (1985). *Acquisition of Hebrew*. Hillsdale: Erlbaum.

Berman, R. A. (1987). Productivity in the lexicon: New-word formation in Modern Hebrew. *Folia Linguistica*, *21*, 425–461.

Berman, R. A. (1990). On acquiring an (S)VO language: Subjectless sentences in children's Hebrew. *Linguistics*, *28*, 1135–1166.

Berman, R. A. (1993a). Marking of verb transitivity by Hebrew-speaking children. *Journal of Child Language*, *20*, 641–669.

Berman, R. A. (1993b). Developmental perspectives on transitivity: A confluence of cues. In Y. Levy (Ed.), *Other children, other languages: Issues in the theory of acquisition* (pp. 189–241). Hillsdale: Erlbaum.

Berman, R. A. (2000). Children's innovative verbs vs. nouns: Structured elicitations and spontaneous coinages. In L. Menn & N. Bernstein-Ratner (Eds.), *Methods for studying language production* (pp. 69–93). Mahwah: Erlbaum.

Berman, R. A. (2003). Children's lexical innovations: developmental perspectives on Hebrew verb-structure. In J. Shimron (Ed.), *Language processing and acquisition in languages of Semitic, root-based morphology* (pp. 243–291). Amsterdam: Benjamins.

Berman, R. A. (2012). Revisiting roots in Hebrew: A multi-faceted view. In M. Muchnik & T. Sadan (Eds.), *Studies on Modern Hebrew and Jewish languages in honor of Ora (Rodriguez) Schwarzwald* (pp. 132–158) Jerusalem: Carmel Press.

Berman, R. A., & Armon-Lotem, S. (1997). How grammatical are early verbs? In C. Martinot (Ed.), *Annales littéraires de l'Université de Franche-Comte: Vol. 631. Actes du colloque international sur l'acquisition de la syntaxe en langue maternelle et en langue etrangère* (pp. 17–59).

Berman, R., & Dromi, E. (1984). On marking time without aspect in child language. *Papers and Reports on Child Language Development*, *23*, 23–32.

Berman, R. A., & Katzenberger, I. (2004). Form and function in introducing narrative and expository texts: A developmental perspective. *Discourse Processes*, *38*, 57–94.

Berman, R. A., & Nir-Sagiv, B. (2004). Linguistic indicators of inter-genre differentiation in later language development. *Journal of Child Language*, *31*, 339–380.

Berman, R. A., & Nir-Sagiv, B. (2007). Comparing narrative and expository text construction across adolescence: A developmental paradox. *Discourse Processes*, *43*, 79–120.

Berman, R. A., & Nir-Sagiv, B. (2010). The lexicon in speech-writing differentiation: Developmental perspectives. *Written Language and Literacy*, *13*, 181–203.

Berman, R. A., & Ravid, D. (2009). Becoming a literate language user: Oral and written text construction across adolescence. In D. R. Olson & N. Torrance (Eds.), *Cambridge handbook of literacy* (pp. 92–111). Cambridge: Cambridge University Press.

Berman, R. A., & Sagi, Y. (1981). Children's word-formation and lexical innovations. *Hebrew Computational Linguistics Bulletin*, *18*, 31–62 [in Hebrew].

Berman, R. A., & Verhoeven, L. (2002). Developing text production abilities in speech and writing: Aims and methodology. *Written Language and Literacy*, *5*, 1–44.

Blevins, J. P. (2014). The morphology of words. In M. Goldrick, V. S. Ferreira, & M. Miozzo (Eds.), *The Oxford handbook of language production*, Oxford: Oxford University Press.

Blevins, J. P. (2016). *Word and paradigm morphology*. Oxford: Oxford University Press.

Bolozky, S. (1997). Israeli Hebrew phonology. *Phonologies of Asia and Africa*, *1*, 287–311.

Bolozky, S. (1999). *Measuring productivity in word formation: The case of Israeli Hebrew*. Leiden: Brill.

Bolozky, S. (2003). The 'roots' of denominative Hebrew verbs. In J. Shimron (Ed.), *Language processing and acquisition in languages of Semitic, root-based morphology* (pp. 131–146). Amsterdam: Benjamins.

Bolozky, S. (2004). Review of Uzzi Ornan, the final word: Mechanism for Hebrew word generation. *Hebrew Studies*, *45*, 285–287.

Bolozky, S. (2007). Israeli Hebrew morphology. In A. S. Kaye (Ed.), *Morphologies of Asia and Africa (including the Caucasus)* (pp. 283–308). Winona Lake: Eisenbrauns.

Bolozky, S. (2009). Frequency and productivity in the verb system of Israeli Hebrew. *Leshonenu*, *71*(2/3), 345–367 [In Hebrew].

Bonami, O., & Stump, G. T. (2016). Paradigm function morphology. In A. Hippisley & G. T. Stump (Eds.), *Cambridge handbook of morphology* (pp. 449–481). Cambridge: Cambridge University Press.

Boudelaa, S., & Marslen-Wilson, W. D. (2005). Discontinuous morphology in time: Incremental masked priming in Arabic. *Language and Cognitive Processes*, *20*, 2017–2260.

Brysbaert, M., Keuleers, E., & New, B. (2011). Assessing the usefulness of Google books' word frequencies for psycholinguistic research on word processing. *Frontiers in Psychology*, *2*, 27.

Bybee, J. L. (1988). Morphology as lexical organization. In M. Hammond & M. Noonan (Eds.), *Theoretical morphology* (pp. 119–141). San Diego: Academic Press.

Clark, E. V. (2003). *The lexicon in acquisition*. Cambridge: Cambridge University Press.

Clark, E., & Clark, H. H. (1979). When nouns surface as verbs. *Language*, *55*, 767–811.

Dattner, E. (2015). Enabling and allowing in Hebrew: A usage-based construction grammar account. In B. Nolan, G. Rawoens, & E. Diedrichsen (Eds.), *Causation, permission, and transfer: Argument realization in GET, TAKE, PUT, GIVE and LET verbs* (pp. 271–293). Amsterdam: Benjamins.

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.

Deutsch, A., & Malinovitch, T. (2016). The role of the morpho-phonological word-pattern unit in single-word production in Hebrew. *Journal of Memory and Language*, *87*, 1–15.

Deutsch, A., & Meir, A. (2011). The role of the root morpheme in mediating word production in Hebrew. *Language and Cognitive Processes*, *26*, 716–744.

Diessel, H., & Hilpert, M. (2016). Frequency effects in grammar. In M. Aronoff (Ed.), *Oxford research encyclopedia of linguistics*, New York: Oxford University Press.

Dressler, W. U. (2005). Word-formation in natural morphology. In P. Stekauer & R. Lieber (Eds.), *Handbook of word-formation* (pp. 267–284). New York: Springer.

Dressler, W. U., Barbaresi, L. M., Schwaiger, S., Ransmayr, J., Sommer-Lolei, S., & Korecky-Kröll, K. (2019). Rivalry and lack of blocking among Italian and German diminutives in adult and child language. In F. Rainer, F. Gardani, H. C. Luschützky, & W. U. Dressler (Eds.), *Competition in inflection and word-formation* (pp. 123–143). Cham: Springer.

Dromi, E. (1987). *Early lexical development*. Cambridge: Cambridge University Press.

Dromi, E., & Berman, R. A. (1986). Language-general and language-specific in developing syntax. *Journal of Child Language*, *14*, 371–387.

Eitan, S. (2015). *The Semitic root as a window to verb acquisition in early childhood*. MA thesis, Department of Communications Disorders, Tel Aviv University.

Elman, J. L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition*, *48*, 71–99.

Ephratt, M. (1997). The psycholinguistic status of the root in Modern Hebrew. *Folia Linguistica*, *31*, 77–103.

Frost, R. (2012). Towards a universal model of reading: Target article. *Behavioral and Brain Sciences*, *35*, 263–279.

Frost, R., Deutsch, A., & Forster, K. I. (2000). Decomposing morphologically complex words in a nonlinear morphology. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *26*, 751–765.

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*, 829–856.

Frost, R., & Plaut, D. (2001). *The word-frequency database for Hebrew*. Retrieved January 30, 2018, from http://word-freq.mscc.huji.ac.il/.

Gillis, S., & Ravid, D. (2006). Typological effects on spelling development: A crosslinguistic study of Hebrew and Dutch. *Journal of Child Language*, *33*, 621–659.

Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. New York: Oxford University Press.

Goodman, J. C., Dale, P. S., & Li, P. (2008). Does frequency count? Parental input and the acquisition of vocabulary. *Journal of Child Language*, *35*, 515–531.

Gray, R. M. (2011). *Entropy and information theory*. New York: Springer.

Grunwald, T. (2014). *Verb root spelling in young readers' texts*. MA thesis, School of Education, Tel Aviv University.

Hare, M., & Elman, J. (1995). Learning and morphological change. *Cognition*, *56*, 61–98.

Haspelmath, M., & Sims, A. (2013). *Understanding morphology* (2nd ed.). London: Hodder Education.

Hirsh-Pasek, K., & Golinkoff, R. M. (Eds.) (2006). *Action meets word: How children learn verbs*. New York: Oxford University Press.

Itai, A., & Wintner, S. (2008). Language resources for Hebrew. *Language Resources and Evaluation*, *42*, 75–98.

Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective of cognitive science*. Cambridge: MIT Press.

Kassovsky, O. (1985). The ability to comprehend denominative verbs. *Hebrew Computational Linguistics*, *23*, 81–92 [in Hebrew].

Kastner, I. (2019). Templatic morphology as an emergent property. *Natural Language and Linguistic Theory*, *37*(2), 571–619.

Keuleers, E., Diependaele, K., & Brysbaert, M. (2010). Practice effects in large-scale visual word recognition studies: A lexical decision study on 14,000 Dutch mono- and disyllabic words and nonwords. *Frontiers in Psychology*, *1*, 174.

Keuleers, E., Stevens, M., Mandera, P., & Brysbaert, M. (2015). Word knowledge in the crowd: Measuring vocabulary size and word prevalence in a massive online experiment. *Quarterly Journal of Experimental Psychology*, *68*, 1665–1692.

Keuleers, E., & Marelli, M. (2020). Resources for mental lexicon research: A delicate ecosystem. In V. Pirrelli, I. Plag, & W. U. Dressler (Eds.), *Word knowledge and word usage: A cross-disciplinary guide to the mental lexicon*. Berlin: De Gruyter.

Kidd, E., Lieven, E. V. M., & Tomasello, M. (2010). Lexical frequency and exemplar-based learning effects in language acquisition: Evidence from sentential complements. *Language Sciences*, *32*, 32–142.

Kuznetsova, J. (2015). *Linguistic profiles: Going from form to meaning via statistics*. Berlin: De Gruyter Mouton.

Laks, L. (2013). Why and how do Hebrew verbs change their form? A morpho-thematic account. *Morphology*, *23*, 351–383.

Laks, L. (2018). Verb innovation in Hebrew and Palestinian Arabic: The interaction of morpho-phonological and thematic-semantic criteria. *Brill's Journal of Afroasiatic Languages and Linguistics*, *10*(2), 238–284.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction and representation of knowledge. *Psychological Review*, *104*, 211–240.

Levie, R. (2012). *Morphological development in the shadow of language impairment and SES background*. Unpublished doctoral dissertation, Tel Aviv University.

Levie, R., Ben Zvi, G., & Ravid, D. (2017). Morpho-lexical development in language-impaired and typically developing Hebrew-speaking children from two SES backgrounds. *Reading & Writing*, *30*, 1035–1064.

Levie, R., Dattner, E., Zwilling, R., Rosenstein, H., Stanzas Eitan, S., & 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.

Linzen, T. l. (2009). *Corpus of blog postings collected from the Israblog website*. Tel Aviv: Tel Aviv University.

Lustigman, L. (2013). Developing structural specification: Productivity in early Hebrew verb usage. *First Language*, *33*, 47–67.

Lustigman, L. (2016). From opacity to transparency: Transitional categories in early Hebrew grammer. In R. Berman (Ed.), *Acquisition and development of Hebrew: From infancy to adolescence* (pp. 225–258). Amsterdam: Benjamins.

Marslen-Wilson, W. D. (2007). Morphological processes in language comprehension. In G. Gaskell (Ed.), *Oxford handbook of psycholinguistics* (pp. 175–193). Oxford: Oxford University Press.

Matthews, D., Lieven, E., Theakston, A., & Tomasello, M. (2005). The role of frequency in the acquisition of English word order. *Cognitive Development*, *20*, 121–136.
Mattiello, E., & Dressler, W. U. (2019). The morphosemantic transparency/opacity of novel English analogical compounds and compound families. *Studia Anglica Posnaniensia*, *53*, 67–114. https://doi.org/10.2478/stap-2018-0004.
McCarthy, J. J. (1981). A prosodic theory of nonconcatenative morphology. *Linguistic Inquiry*, *12*, 373–418.
McCauley, S. M., & Christiansen, M. H. (2019). Language learning as language use: A cross-linguistic model of child language development. *Psychological Review*, *126*, 1–51.
Nicoladis, E., Palmer, A., & Marentette, P. (2007). The role of type and token frequency in using past tense morphemes correctly. *Developmental Science*, *10*, 237–254.
Ninio, A. (1999). Pathbreaking verbs in syntactic development and the question of prototypical transitivity. *Journal of Child Language*, *26*, 619–653.
Nir, R. (1993). *Lexical devices in Modern Hebrew*. Tel Aviv: The Open University of Israel [in Hebrew].
Ornan, U. (2014). Innovation of verbs in Hebrew. In *Proceedings of the joint workshop on social dynamics and personal attributes in social media* (pp. 94–97). Baltimore: Association for Computational Linguistics.
Paterson, K. B., Alcock, A., & Liversedge, S. P. (2011). Morphological priming during reading: Evidence from eye movements. *Language and Cognitive Processes*, *26*, 600–623.
Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*, 357–383.
Ramscar, M., Dye, M., & Klein, J. (2013). Children value informativity over logic in word learning. *Psychological Science*, *24*, 1017–1023.
Ravid, D. (1990). Internal structure constraints on new-word formation devices in Modern Hebrew. *Folia Linguistica*, *24*, 289–346.
Ravid, D. (1995). *Language change in child and adult Hebrew: A psycholinguistic perspective*. New York: Oxford University Press.
Ravid, D. (2001). Learning to spell in Hebrew: Phonological and morphological factors. *Reading & Writing*, *14*, 459–485.
Ravid, D. (2003). A developmental perspective on root perception in Hebrew and Palestinian Arabic. In Y. Shimron (Ed.), *Language processing and acquisition in languages of Semitic, root-based morphology* (pp. 293–319). Amsterdam: Benjamins.
Ravid, D. (2004). Later lexical development in Hebrew: derivational morphology revisited. In R. A. Berman (Ed.), *Language development across childhood and adolescence: Psycholinguistic and crosslinguistic perspectives* (pp. 53–82). Amsterdam: Benjamins.
Ravid, D. (2006). Word-level morphology: A psycholinguistic perspective on linear formation in Hebrew nominals. *Morphology*, *16*, 127–148.
Ravid, D. (2012). *Spelling morphology: the psycholinguistics of Hebrew spelling*. New York: Springer.
Ravid, D. (2019). Derivation. In R. A. Berman (Ed.), *Usage-based studies in Modern Hebrew: Background, morpho-lexicon, and syntax* (pp. 203–265). Amsterdam: Benjamins.
Ravid, D., Ashkenazi, O., Levie, R., Ben Zadok, G., Grunwald, T., Bratslavsky, R., & Gillis, S. (2016). Foundations of the early root category: Analyses of linguistic input to Hebrew-speaking children. In R. A. Berman (Ed.), *Acquisition and development of Hebrew: From infancy to adolescence* (pp. 95–134). Amsterdam: Benjamins.
Ravid, D., & Bar On, A. (2005). Manipulating written Hebrew roots across development: The interface of semantic, phonological and orthographic factors. *Reading & Writing*, *18*, 231–256.
Ravid, D., & Hershkovitz, L. (2017). The development of Hebrew conjunct constructions in narration. In E. Segers & P. van den Broek (Eds.), *Developmental perspectives in written language and literacy* (pp. 119–135). Amsterdam: Benjamins.
Ravid, D., Levie, R., & Avivi Ben Zvi, G. (2003). The role of language typology in linguistic development: Implications for the study of language disorders. In Y. Levy & J. Schaeffer (Eds.), *Language competence across populations: Toward a definition of SLI* (pp. 183–207). Mahwah: Erlbaum.
Ravid, D., & Malenky, D. (2001). Awareness of linear and nonlinear morphology in Hebrew: A developmental study. *First Language*, *21*, 25–56.
Ravid, D., & Schiff, R. (2006a). Morphological abilities in Hebrew-speaking gradeschoolers from two socio-economic backgrounds: An analogy task. *First Language*, *26*, 381–402.
Ravid, D., & Schiff, R. (2006b). Roots and patterns in Hebrew language development: Evidence from written morphological analogies. *Reading & Writing*, *19*, 789–818.

Ravid, D., & Schiff, R. (2012). From dichotomy to divergence: Number/gender marking on Hebrew nouns and adjectives across schoolage. *Language Learning*, *62*, 133–169.

Ravid, D., & Vered, L. (2017). Hebrew verbal passives in later language development: The interface of register and verb morphology. *Journal of Child Language*, *44*, 1309–1336.

Schiff, R., & Ravid, D. (2007). Morphological analogies in Hebrew-speaking university students with dyslexia compared with typically developing gradeschoolers. *Journal of Psycholinguistic Research*, *36*, 237–253.

Schwarzwald, O. R. (1981). *Grammar and reality in the Hebrew verb*. Ramat Gan: Bar Ilan University Press.

Schwarzwald, O. R. (1996). Syllable structure, alternations and verb complexity: The Modern Hebrew verb patterns re-examined. *Israel Oriental Studies*, *16*, 99–112.

Schwarzwald, O. R. (2000). Verbal roots and their links to nouns. In O. R. Schwarzwald, S. Blum-Kulka, & E. Olshtain (Eds.), *Studies in the media, linguistics and language teaching—Raphael Nir jubilee book* (pp. 426–438). Jerusalem: Carmel [in Hebrew].

Schwarzwald, O. R. (2001). Derivation and innovation in Hebrew: Quantitative aspects. In O. R. Schwarzwald & R. Nir (Eds.), *Studies in Hebrew and language teaching in honor of Ben Zion Fischler* (pp. 265–275). Even Yehuda: Reches [in Hebrew].

Schwarzwald, O. R. (2002). *Studies in Hebrew morphology*. Tel Aviv: The Open University of Israel Press [in Hebrew].

Schwarzwald, O. R. (2013). *Defective verbs (I: 673-678). EHLL: Encyclopedia of Hebrew Language and Linguistics, I–IV volumes*. New York: Brill.

Schwarzwald, O. R., & Neradim, E. (1995). Hebrew *šaf'el*. *Lešonenu*, *58*, 145–152 [in Hebrew].

Seroussi, B. (2011). *The morphology-semantics interface in the mental lexicon: the case of Hebrew*. Unpublished doctoral dissertation, Tel Aviv University.

Sivan, R. (1976). *On the foundations of present-day Hebrew*. Tel Aviv: Rubinshtein [in Hebrew].

Temkin Martínez, M. (2010). *Sources of non-conformity in phonology: Variation and exceptionality in Modern Hebrew spirantization*. Ph.D. dissertation, University of Southern California.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MASS: Harvard University Press.

Traugott, E. C., & Trousdale, G. (2013). *Constructionalization and constructional changes*. Oxford: Oxford University Press.

Ussishkin, A. (2005). A fixed prosodic theory of nonconcatenative templatic morphology. *Natural Language and Linguistic Theory*, *23*, 169–218.

Velan, H., Frost, R., Deutsch, A., & Plaut, D. C. (2005). The processing of root morphemes in Hebrew: Contrasting localist and distributed accounts. *Language and Cognitive Processes*, *20*, 169–206.

Vermunt, J. K., & Magidson, J. (2002). Latent class cluster analysis. *Applied Latent Class Analysis*, *11*, 89–106.

Zwilling, R. (2009). *Noun plurals in children's triadic peer talk 2–8 years*. MA thesis, Department of Communications Disorders, Tel Aviv University.