Breaking into the Hebrew verb system: a learning problem

Keywords

Verb acquisition, Hebrew, Child Directed Speech, early child language, inflectional morphology, tense and modality

Abstract

Verb learning is an important part of linguistic acquisition. The present study examines the early phases of verb acquisition in Hebrew, a language with complex derivational and inflectional verb morphology, analyzing verbs in dense recordings of CDS and CS of two Hebrew-speaking parent-child dyads aged 1;8-2;2. Our goal was to pinpoint those cues that help toddlers identify the root-and-pattern make-up of Hebrew verbs despite the prevalence of structural opacity and irregularity in the verbs they hear, due to a high token frequency of defective (irregular)-root verbs. The study provides a detailed account of the distribution of root types and temporal categories in Hebrew CDS and CS showing how verb specific morphological features in the form of inflectional affixes in the Modal Cluster of Infinitive Imperative and Future Tense in CDS act as distributional cues facilitating verb acquisition in CS.

Introduction

The verb is a central lexical category of human languages (Schachter, 1985), expressing the relationships between objects and individuals and collective terms. Carrying morpho-syntactic, lexico-semantic and temporal information, verbs constitute the “architectural centerpiece” of grammar, as they determine the argument structure of the sentence (Golinkoff & Hirsh Pasek, 2006:4). As relational terms labeling actions, events, processes and states, verbs are harder to conceptualize by children than concrete nouns (Gentner, 1982; Golinkoff & Hirsh-Pasek, 2008). Nonetheless, in most of the world’s languages, acquisition of core verb inflection is achieved by three years of age, regardless of the complexity of the morphological system of the target language (Bittner, Dressler & Kilani-Schoch, 2003; Tomasello, 2006). Children approach verb learning in ways that are specific to their native language, given the differential typological organization of verb morphology and lexical semantics (Kibrik, 2012; Koptjevskaja-Tamm, 2012; Talmy, 1985). The current study examines the information that makes it possible for Hebrew-speaking toddlers to begin deconstructing and constructing verb morphology.

Scholars working within Usage Based acquisition approaches consider Child Directed Speech (CDS) as a major domain where statistical learning – tracking predictive relationships between elemental units (Yurovsky, Yu & Smith, 2012) or extracting linguistic structures based on regularities in the input (Romberg & Saffran, 2010) - takes place. This is the arena where words and morphemes are extracted from the speech stream, patterns are detected, and children gain linguistic generalizations, helped by socio-cognitive abilities (Saffran, 2003; Tomasello, 2003, 2006, 2009). Schemas and abstract categories gradually emerge out of the items children have learned, based on the distributional and frequency properties of the input (Behrens, 2006; Lieven, 2008; Lieven, Behrens, Spears & Tomasello, 2003; Tomasello, 2004, 2006).

Verb learning is an important part of linguistic acquisition. Children as young as 13.5 months rely on prosodic and phonological cues to detect verbs in the ambient language (Golinkoff & Hirsch-Pasek, 2006; Nazzi & Houston, 2006). One such distributional cue is frequent frames. These may include verbs and play an important role in carrying out the verb learning task, constituting a form of non-adjacent dependency to which infants are sensitive, making them a salient distributional environment for young language learners (Gómez, 2002, 2008). Several types of frequent frames were described in the literature. One of them is relevant to languages with sparse morphology, such as English, where verbs are signaled by planting them within different two-word utterances, e.g., *he___ it*, *who ___ some* (Mintz, 2006). 12-month old infants demonstrate special sensitivity to verbs in these frames (this does not work for German and Dutch, see Stumper, Bannard, Lieven & Tomasello, 2011). Another type defined by Sandoval & Gómez (2013) as 'morpheme frames', is found in heavily inflected languages such as Hebrew, highlighting verb structure via a limited number of extremely frequent inflectional morphemes at the edges of the verb (Mintz, 2003).

Against this background, the present study examines the early phases of verb acquisition in Hebrew, a language with complex derivational and inflectional verb morphology (Ravid, 2012), aiming to pinpoint those cues that help toddlers identify the makeup of Hebrew verbs.

Hebrew verb structure

Hebrew is a Semitic language with rich morphology which organizes its entire lexicon (Berman, 1987; Schwarzwald, 2002). A fundamental structure in Semitic languages is the non-linear combination of non-continuous consonantal roots and (mainly) vocalic patterns (Bolozky, 2007; Ravid, 2012), as in *gadal* ‘grow’ and *migdal* ‘tower’, which share root *g-d-l* and differ in the affixed pattern. Verbs are always derived non-linearly by combining a root with one of five[1] verb patterns termed *binyanim* (literally ‘buildings’) – traditionally named *Qal, Nif'al, Hif'il, Pi'el,* and *Hitpa'el*. Root-based verbs with different *binyan* patterns constitute derivational families with members expressing different transitivity and Aktiosart values (Bolozky, 2007; Ravid, 2003; Schwarzwald, 2002). For example, root *l-m-d* 'learn' (in bold) combines with *binyan* derivational patterns to create a family of four different verbs: basic ***lamad*** 'learn', passive *ni*__*lmad*__ 'be learned', causative ***limed*** 'teach', and middle *hit*__*lamed*__ 'apprentice' (Berman, 1975; Schwarzwald, 1998).

Root and pattern nonlinear combination serves verb structure in yet another way, which is the focus of the current study. What is traditionally known as '*binyan*' is not a single pattern, but actually a bundle of temporal patterns which combine with the same root to construct a set of temporal stems unique to each *binyan*. For example, *CaCaC, CoCeC* and *li-CCoC* serve as the respective past, present and infinitive patterns of *Qal*, with the capital *Cs* representing root radicals. Thus the shift from past tense ***lamad*** 'learned' to present tense ***lomed*** 'learns' and infinitive *li-*__*lmod*__ 'to-learn' involves changing the temporal patterns across the same root *l-m-d.* Table 1 presents the five non-passive *binyanim* as sets of temporal stems (root radicals appear as uppercase C's; P represents temporal/agreement prefixes in future tense).

Table 1. The temporal stems of the five non-passive *binyanim*

***Binyan*/temporal stem**	**Infinitive**	**Imperative**	**Future**	**Present**	**Past**
Qal	*liCCoC*	*tiCCoC*	*PiCCoC*	*CoCeC*	*CaCaC*
Nif'al	*lehiCaCeC*	*tiCaCeC*	*PiCaCeC*	*niCCaC*	*niCCaC*
Hif'il	*lehaCCiC*	*taCCiC*	*PaCCiC*	*maCCiC*	*hiCCiC*
Pi'el	*leCaCeC*	*teCaCeC*	*PeCaCeC*	*meCaCeC*	*CiCeC*
Hitpa'el	*lehitCaCeC*	*titCaCeC*	*PitCaCeC*	*mitCaCeC*	*hitCaCeC*

Hebrew verb acquisition

Studies on Hebrew verb acquisition in the second year of life report that child speech contains verbs mostly in the basic *Qal* pattern, with restricted temporal expression, mostly consisting of modal, imperative or infinitive forms, or present-tense modal verbs such as *rotse* 'wants' and *yaxol* 'able' (Berman & Armon Lotem, 1996). A small number of finite verbs in this age range were found to encode mainly aspectual values, such as Present Tense for continuity, e.g., *holex* 'is walking', and Past Tense for punctuality and change of state, such as *nafal* 'fell', *nigmar* 'all gone' (Berman, 1985a,b; Ravid, 1997). Early Hebrew verb structure was found to be restricted to a single inflectional form (e.g., past first person singular), with many opaque, so-called ‘bare stems’ that could stand for several tense-mood categories. For example, the truncated form *pes* could be interpreted *le-tapes* 'to-climb', *metapes* 'is climbing' and *tipes* 'climbed’ (Armon Lotem & Berman, 2003).

According to Berman & Armon-Lotem (1996, 2003), the growing ability to inflect verbs in Hebrew is expressed in a developmental decrease in 'bare-stem' forms, and a transition from formulaic rote-learned forms to a rule-based system, leading to mastery of the major inflectional categories of the verb by age 3. A recent study by Lustigman (2012, 2013) investigated the development of verb inflections in four toddlers aged 1;4-2;2 in the context of Berman's phase model (1986). According to

Lustigman, Hebrew speaking children initially use bare stems together with rote-learned verbs, while prefix-marked full infinitives occur later, in well-formed syntactic contexts, alongside the productive use of affixed present tense participle forms. Finally, Both Kaplan (1983) and Herzberg (2010) noted that Past Tense agreement in second person as the last and most difficult person marking to appear.

The next section presents further analysis of the morphological components of Hebrew verbs, leading to a suggested path for Hebrew verb acquisition.

Inflectional verb morphology: The early challenge

The pervasive non-linear form of Hebrew verbs means that Hebrew speaking children need to gain command of root and pattern structure not only for verb derivation (Berman, 1993a, b), but first and foremost for temporal inflection. That is, they need to learn the specific set of temporal patterns constructing each *binyan* conjugation (Berman, 1982; Ravid, 2012). It stands to reason that learning the structures expressing temporality is an important initial step in Hebrew verb acquisition (Berman & Slobin, 1994; Lustigman, 2012), laying the foundations of verb structure. This job looks doable assuming the consistent re-occurrence of root radicals within the changing vocalic templates of temporal patterns, which facilitates the perception of root and pattern structure (Gómez, 2008). When roots are full (regular), as illustrated in Table 1, stem morpho-phonology is indeed transparent. However, when temporal stem form is based on defective (irregular) roots, stem morpho-phonology is rendered opaque and less discernible (Ravid, 1995), as described below, making the job of root and pattern learning challenging.

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 (sometimes four) consonantal root radicals constructing canonical, transparent stems where root and pattern structure can be easily identified. Such verb structures, based on full roots, are illustrated by root *l-m-d* above and the canonical temporal stems in Table 1. Defective roots may be considered as the irregular Hebrew root category. They mostly[2] contain non-consonantal, weak radicals such as *y,* w or *ʔ,* yielding non-canonical, opaque stems (Berman, 2003; Ravid, 1995, 2012).

Stems based on full, regular roots are transparent in two senses – all root radicals always show up in the stem as a set of easily identifiable consonants (e.g., *l-m-d* ‘learn’); this repeating consonantal set makes it easy to discern the changing vocalic pattern of temporal stems (Table 1). Full roots thus optimize learning of the root-and-pattern non-linear structure of Hebrew verbs. This is illustrated by the following examples of infinitive *Qal* based on full roots: *li-shmor* ('to keep', *š-m-r*), *li-sgor* ('to close' *s-g-r*), *li-lmod* (to study*, l-m-d*) and *li-xtov* ('to write', *k-t-b*) clearly sharing the same pattern *li-CCoC.*

In contrast, stems based on defective roots are opaque in the sense of often containing only a part of the root (e.g., only the *b* of root *b-w-ʔ* ‘come’, often showing up as *v* (Table 3). Moreover, defective stems are phonologically variant and fused, making it difficult to identify the root and pattern components. This is illustrated by the following examples of *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*), and *la-tset* ('to-go out', *y-c-ʔ*). These verbs refer to familiar, salient activities, making them highly frequent in young children's lexicons, and consequently resistant to regularizing change (Berman & Armon Lotem, 1986, 2003, Dromi, 1987; Hare & Elman, 1995).

Stems based on defective roots thus pose a learning challenge to the Hebrew-acquiring toddler as they make it difficult to perceive the non-linear structure central to Hebrew temporal inflection and to verb derivation. Given the high prevalence of irregular forms in spoken language in general (Bybee, 2006), and in CS and CDS in particular (Davis, 2009; Rose, Stevenson & Whitehead, 2002), we can assume that such opaque forms constitute a real obstacle to learning root and pattern structure in Hebrew.

The proposed model - agreement boundaries as anchor cues

Assuming that stems encountered in early Hebrew CDS are often opaque and variegated, based on irregular roots (Ravid, 2010), we propose that young Hebrew learners initially turn to the structurally clear, stable, salient affixal boundaries at stem edges (Gómez, 2002, 2008; Mintz, 2003, 2006; Ravid, 2012). Table 1 shows, for example, that the infinitival *l-* 'to' is a stable cue across all *binyanim*. In the same way, *m-* and *h-* denote Present and Past Tense respectively in several *binyanim*. These boundaries also serve to mark agreement with the grammatical subject in gender, number and person[3]. Agreement is marked uniformly across all *binyanim* by prefixes and suffixes attached to the temporal stems, as shown in Table 2. For example, *telxu* 'You,Pl will go' (based on defective root *y-l-k* 'go' in *Qal*) is framed by prefix *t-* marking second person in future / imperative and suffix *–u* denoting plural number. Affixal cues are particularly salient in the three modal forms - the infinitive, imperative and future tense forms[4], labeled together 'the modal cluster' (Ashkenazi, 2015; Ravid, 2010)[5]: This is the only temporal verb site where markers can occur as both prefixes and suffixes across all *binyanim* (Table 2).

Table 2. Agreement markers on Hebrew verbs

Temporal category	prefix		Suffix	Person	Number	Gender
Imperative	-	TEMPORAL STEM	*-i*	2nd	Singular	Feminine
	-		∅	2nd	Singular	Masculine
	-		*-u*	2nd	Plural	-
Future	*ʔ*		-	1st	Singular	-
	t		-	2nd /3rd	Singular	Masculine/feminine
	t		*-i*	2nd	Singular	Feminine
	t		*-u*	2nd	Plural	-
	P		∅	3rd	Singular	Masculine
	P		*-u*	3rd	Plural	-
	n		-	1st	Plural	-
Past	-		*-ti*	1st	Singular	-
	-		*-ta*	2nd	Singular	Masculine
	-		*-t*	2nd	Singular	Feminine
	-		∅	3rd	Singular	Masculine
	-		*-a*	3rd	Singular	Feminine
	-		*-nu*	1st	Plural	-
	-		*-tem*	2nd	Plural	Masculine
	-		*-ten*	2nd	Plural	Feminine
	-		*-u*	3rd	Plural	-
Present	-		∅	-	Singular	Masculine
	-		*-a/-et*	-	Singular	Feminine
	-		*-im*	-	Plural	Masculine
	-		*-ot*	-	Plural	Feminine
	-		∅	-	Singular	Masculine

Against this background, the current study proposes a new account of the early learning of the Hebrew verb system which differs from the previous investigations of early Hebrew verb acquisition in two important ways. First, it focuses on the root-and-pattern verb composition that has mainly been associated with later, derivational morphology development, showing why this structure is critical for the early learning of inflectional verb morphology. Second, it incorporates information on the development of verb structure in both CDS and CS, thus highlighting the foundations of verb learning in both parental input and child output and the relations between them. The issue of input - output relations in the acquisition of verb morphology has been investigated in other languages such as Turkish (Aksu-Koc, 1998), French

(Veneziano & Parisse, 2010; Clark & De Marneffe, 2010), and a group of nine typologically different languages including Dutch, German, Russian, Greek and Yukatan Maya (Xanthos et al., 2011). It was studied briefly in Hebrew relating to the use of past tense verbs (Ninio, 2008) and the use of transitive verbs in single word utterances (Ninio, 2015). However, no study to date has investigated the role of CDS in Hebrew verb acquisition focusing on its structural makeup.

The proposed model presents an account of the early learning of the Hebrew verb system, with modal cluster stems based on defective roots, especially in *Qal*, serving as its springboard. Our hypothesis relies crucially on a three-pronged prediction regarding the initial distributions of verbform tokens in Hebrew CDS and CS. First, Temporality-wise, most such tokens will initially consist of the modal cluster, serving basic pragmatic content of early caregiver-child interaction (Aikhenvald, 2010; Berman, 1985b; Stephany, 1983). Second, structurally, they will mostly be composed of defective roots as the irregular verb inventory prevalent in spoken and child-oriented speech in particular (Berman & Armon Lotem, 1996, 2003; Dromi, 1987) and as token frequency of irregular forms is high in language in general (Bybee, 2006). And third, *Binyan*-wise, they will mostly occur in *Qal* (Berman, 1993a, b), the verb pattern with the highest frequency in all Hebrew corpora (Berman, Nayditz & Ravid, 2011). This combination creates particularly opaque stems, on the one hand, with particularly salient, stable boundaries (agreement markers), on the other (Table 2).

Method

The current analysis focuses on type and token frequencies of structural root categories and temporal stems in input-output databases of two Hebrew-speaking parent-child dyads.

Participants

The analyses reported below are based on a densely recorded corpus of naturalistic longitudinal interactions of two Hebrew-speaking parent-child dyads - a boy dyad and a girl dyad. The boy dyad was recorded between the ages 1;08.27 (1 year 8 months and 27 days, or 635 days) to 2;2.03 (2 years 2 months and 3 days, or 795 days), yielding 49 recording sessions. The girl dyad was recorded between the ages of 1;09.25 (1 year 9 months and 25 days, or 664 days) to 2;02.19 (2 years 2 months and 19 days, or 810 days), yielding 47 recording sessions. Different child genders were chosen so as to permit analysis of the obligatory gender agreement in Hebrew verb inflection (Ravid & Schiff, 2015; Schwarzwald, 1998). Both families were from mid-high SES background, living in central Israel. The two sets of parents, who did not know each other, were native speakers of Hebrew, born in Israel and spoke only Hebrew at home. They did not receive any monetary remuneration for their voluntary participation.

Both children were first-born and had no siblings at the time of recording. Both had normal cognitive, communicative and linguistic development according to parental report (including the Hebrew CDI checklist in Maital, Dromi, Sagi & Borenstein, 2000), period assessment at the local neonate and children's health clinic, and assessment by the first author (a certified senior SLP). Neither of them had a history of ear infections or any other major health issues. The boy attended nursery school and the girl did not.

Data collection

The children were audio-recorded by their parents at home during bath time, play time and meal time using an MP3 recorder supplied to the family by the first author. Each dyad was recorded three times a week, for 45-60 minutes each time, for 6 months between 1;8-2;2 approximately (see details above). The parents were informed that the study concerned early language development in Hebrew. They were asked to record spontaneous, natural interactions. Recordings of both dyads started when each child started producing two word utterances and some verbs, based on parental reports using the Hebrew CDI (Maital et al., 2000). Transcriptions of the recordings (see below) ceased when each child produced subject-verb agreement in number and gender in two subsequent recordings, including two different person agreements in past tense, on at least two different verbs. This morpho-syntactic criterion indicated that the child was gaining command of the basic components of verb structure and semantics by productively using temporal stems, which involves root-pattern alternations, as well as agreement markers (Lustigman, 2012; Ravid et al., in press).

Transcription

Dyadic interactions were transcribed in broad phonemic transcription following the CHILDES conventions (MacWhinney, 2005), adapted to take into account Hebrew-specific phonemic, phonological, prosodic and orthographic features (Albert, MacWhinney, Nir & Wintner, 2012). The transcriptions were carried out by undergraduate students of an academic SLP program who took a CHILDES course as part of their studies. The recordings were thoroughly checked by the first author and corrected when necessary, with an estimated 5% error rate. Next Hebrew MOR was run over the transcripts. The verb forms that were not analyzed by the program were identified and manually coded.

Coding

Verbform coding. Verbs in CDS and CS were represented by a code line including information on all morphological components in the following order: root, *binyan*, temporal category (tense or mood), person, gender and number, inflectional prefixes or suffixes, and gloss. Each such unique combination was considered a separate verbform type.

Root coding. Roots were coded for 13 structural root categories (full and defective), as delineated in Table 3.

Table 3. Structural root categories

Category name	Root type	Example
Full	Tri-consonantal roots	*g-d-l* 'grow'
Quadriliteral (three types)	Quadri-consonantal roots	
1. Quadriliteral Reduplicated	Quadri-consonantal roots composed of a reduplicated set	*q-l-q-l* 'spoil'
2. Quadriliteral Final Doubled	Quadri-consonantal roots with the final consonant doubled	*Ç-r-b-b* 'mix'
3. quadriliteral Denominal	Quadri-consonantal roots derived from nominals	*Ç-n-y-n* 'interest' from noun *inyan* 'interest'
Defective Double	Tri-consonantal roots with two identical final consonants, creating non-canonical and opaque morpho-phonological structures	*s-b-b* 'turn around'

Defective *n*-initial	with initial n, which Roots does not appear in all verb forms (deleted in consonant clusters)	*n-p-l*	'fall'
Defective ʔ-initial	Roots with initial ʔ- (A), not pronounced in all verb forms	*a-h-b*	'love'
Defective *y*-initial	Roots with initial y, not pronounced/does not appear in all verb forms	*y-r-d*	'get down'
Defective *y* -medial	with medial y creating non-canonical and opaque morpho-phonological structures	*r-y-b*	'fight'
Defective *w*-medial	with medial w, creating non-canonical and opaque morpho-phonological structures	*q-w-m*	'get up'
Defective ʔ-final	Roots with final ʔ, not pronounced in all verb froms	*m-c-ʔ*	'find'
Defective *y*-final	Roots with final y, which does not appear in all verb forms	*b-k-y*	'cry'
Defective Composites	Roots belonging to more than one defective category	*b-w-ʔ* medial and ʔ-final	'come' (w-

Coding temporal categories. Temporal stems were sorted into four categories, as follows: (1) The Modal Cluster, consisting of the Infinitive, Imperative and Future Tense stems; (2) Present Tense stems; (3) Past Tense stems; and (4) Undefined stems, relevant only in CS. The Undefined category contained two types of non-classifiable forms. One was truncated verb forms, which could phonetically stand for several forms in the syntactic context produced (Lustigman, 2012)[6]. For example, *gid* 'say' in *gid lo* 'say to him', which could mean *tagid* 'Future, 2nd[7], *lehagid* 'Inf.', *nagid* 'Future,1st PL', *agid* 'Future,1st, SG', or *yagid* 'Future, 3rd, MS, SG'. The second non-classifiable form consisted of phonologically unclear verbs, e.g. *tasik* in the phrase *ima tasik iti* 'mommy ? with me'.

Results

We first present the quantitative structure of the database (Table 4), followed by analyses of *binyan*, structural root categories and temporal categories. Table 4 presents the quantitative structure of the CDS and CS databases in terms of words, verbs, and roots.

Table 4. The database

	CDS to BOY	**CDS to GIRL**	**BOY CS**	**GIRL CS**
# of words	140,782	158,679	32,369	39,717
# of Root type	426	438	159	182
# of verb tokens	23,830	31,283	3101	4610

Analyses of type and token distributions of *binyan* , root categories and temporal categories were carried out across the CDS and CS databases for each recording session of each of the two dyads. All statistical analyses were carried out using the

JMP 12 PRO software program. In case of multiple comparisons, this program applies ample Bonferroni adjustments.

Binyan analysis

Table 5 presents the distributions of *binyan* patterns for the CDS and CS lemma types and tokens. It shows that lemma type and token distributions across both dyads were very similar, with *Qal* dominating, followed by *Hif'il*, *Pi'el*, *Nif'al*, or *Hitpa'el* (excluding about 5 tokens of passive *Huf'al* in both CDS data).

Table 5. *Binyan* type and token distributions across CS and CS of both dyads.

	Qal		***Hif'il***		***Pi'el***		***Nif'al***		***Hitpa'el***	
	#	%	#	%	#	%	#	%	#	%
Boy CDS types	5940	61.3	1774	18.3	1262	13.1	300	3.1	407	4.2
Boy CDS tokens	17,327	73.2	3524	14.9	1929	8.1	467	2	435	1.8
Boy CS types	886	76	124	10.6	107	9.2	21	1.8	28	2.4
Boy CS tokens	2684	89.1	145	4.81	113	3.75	31	1.04	39	1.3
Girl CDS types	6532	59.86	1993	18.26	1483	13.58	454	4.16	451	4.13
Girl CDS tokens	22142	70.8	4513	14.4	2873	9.2	805	2.6	945	3
Girl CS types	1337	69.5	319	16.58	129	6.7	82	4.25	57	2.96
Girl CS tokens	3527	76.62	636	13.81	208	4.52	140	3.04	92	2

Root analyses

For the current purposes, the boy and girl databases (CS and CDS) were treated as a single corpus, given the very high correlations in root categories that emerged not only between each child and his or her parents, but also between each parent and the other child, between the two sets of parents, and between the two children (Ashkenazi, Ravid & Gillis, in prep).

Root type frequency was calculated as the sum of the number of single occurrences of roots in each structural category per sesssion. For example, three occurrences of full roots in CDS (e.g., *s-g-r* 'close', *p-t-ħ '*open' and *k-t-b* 'write') in one session would yield a type frequency of three for the category of full roots in that session. Token frequency was calculated as the sum of all occurrences of roots in each structural category per sesssion. For example, if the full roots *s-g-r*, *p-t-ħ,* and *k-t-b* had five, seven, and three occurrences respectively in a certain sesssion, then the token frequency value of the category of full roots for that session would be 15 (5+7+3). Table 6 presents type and token frequencies of the root categories across the CDS and CS databases.

Table 6. Distribution of root categories (raw frequencies and percentages) in CDS and CS of the entire corpus

	CDS		**CS**	
	#	%	#	%
Full root types	379	72.7%	142	64.2%
Defective root	142	27.3%	79	35.8%

types				
Full root tokens	15061	27.5%	1829	23.7%
Defective root tokens	39749	72.5%	5877	76.3%

Type frequency was higher in full roots (tri- and quadri-literal full roots) than in defective roots across CDS and CS; token frequency was higher in defective roots than in full roots across CDS and CS. To further investigate the distribution of full versus defective roots, we analyzed their token frequency in the 20 most common roots in the CS and CDS databases (Table 7). 18 out of the 20 most common roots in CDS (90%) and 17 out of 20 in CS (85%) are defective, with the token frequency of these roots constituting about 60% and 70% of the entire root (which is actually verb) token frequency of CDS and CS respectively, probably due to a high repetition rate of verbs carrying these roots.

Table 7. Token frequencies of the 20 most common roots in CDS and CS

CDS roots	**Token frequency**	**CS roots**	**Token frequency**
b-w-ʔ 'come, bring'	6949	*r-ʔ-y* 'see, look'	1252
r-c-y 'want'	3876	*r-c-y* 'want'	788
r-ʔ-y 'see, look'	3676	*b-w-ʔ* 'come, bring'	638
ʕ-s-y 'do'	2949	*y-š-n* 'sleep'	329
s-y-m 'put'	2406	*ʕ-s-y* 'do'	225
y-l-k 'go, walk'	1458	*p-t-x* 'open'	213
ʔ-m-r 'say'	1256	*l-q-x* 'take'	203
ʔ-k-l 'eat'	1252	*s-y-m* 'put'	191

y-š-b 'sit'	1166	*n-p-l* 'fall'	187
n-g-d 'say'	1102	*y-c-ʔ* 'go out'	166
y-d-ʕ 'know'	971	*ʕ-z-r* 'help'	154
l-q-x 'take'	853	*ʔ-k-l* 'eat'	149
q-r-ʔ 'read'	768	*b-k-y* 'cry'	149
n-t-n 'give'	705	*y-l-k* 'go, walk'	123
y-c-ʔ 'go out'	664	*q-w-m* 'get up'	120
y-š-n 'sleep'	651	*y-š-b* 'sit'	114
s-x-q 'play'	591	*k-ʔ-b* 'hurt'	108
y-k-l 'can'	526	*q-r-ʔ* 'read'	86
n-p-l 'fall'	517	*y-r-d* 'go down'	86
s-k-l 'look'	511	*n-t-n* 'give'	72

Temporal category analyses

Analyses of the temporal categories were carried out separately for the boy dyad and the girl dyad.

Types and tokens of the three temporal categories – the Modal Cluster (Infinitive, Imperative and Future Tense), Present Tense and Past Tense in CDS and CS, and of the Undefined category only in CS, were calculated for each day of recording. A temporal type was defined as a wordform in each of the five temporal stems. For example, Past Tense wordforms *katav* 'wrote, 3rd MS, SG', *lamdu* 'learned, 3rd,PL' and *dafka* 'knocked, 3rd, FM,SG' were each considered a temporal type, hence yielding a type frequency of 3 for Past Tense. Every occurrence of a type was counted as a token. Consider the following example: Three types in Future Tense / Imperative - *tikansi* 'enter,2nd,FM,SG', *tetsayer* 'draw,2nd, MS,SG/3rd Fm,SG', and *nelex* 'go,1st, PL', each occurring three times, yielded a token frequency of 9 for Future Tense[8]. Each undefined verb was examined separately for its possible temporal meaning in context, revealing that most of these verbs were modal in nature. The raw type and token frequencies and the percentages of the temporal categories in CDS and CS of each dyad are presented in Tables 8-11, showing that the modal cluster of Infinitive, Imperative and Future tense forms in CDS and CS (and the undefined with modal meaning only in CS) constitute about 50% of the total verb type and token frequency in both dyads, followed by Present and finally Past tense.

Table 8. Type frequencies of the temporal categories in CDS to boy and girl

	CDS to boy		**CDS to girl**	
	N	**%**	**N**	**%**
Future	2751	28.3	3784	34.6
Infinitive	1458	15.02	1431	13.1
Imperative	380	3.9	564	5.1
Modal cluster	4589	47.2	5779	52.8
Past	2020	20.8	2268	20.7
Present	3098	31.9	2876	26.3

Table 9. Token frequency and percentages of the temporal stems in CDS to boy and girl

	CDS to boy		**CDS to girl**	
	N	**%**	**N**	**%**
Future	6255	26.2	10515	33.6
Infinitive	3188	13.37	3146	10
Imperative	2578	10.8	4903	15.7
Modal cluster	12021	50.37	18564	59.3
Past	3491	14.64	4191	13.4
Present	8325	34.9	8534	27.3

Table 10. Type frequency and percentages of the temporal stems in boy and girl CS

	Boy CS		Girl CS	
	N	**%**	**N**	**%**
Future	171	14.5	344	18
Infinitive	246	20.8	219	11.4
Imperative	129	10.9	216	11.29
Undefined modal	134	11.3	218	11.4
Modal cluster	680	57.5	997	51.83
Past	162	14	390	20.39
Present	324	27.5	495	25.88
Undefined present/past	12	1	30	1.5

Table 11. Token frequency and percentages of temporal stems in boy and girl CS (excluding modal *want* and *can*)

	Boy CS		Girl CS	
	N	**%**	**N**	**%**
Future	278	8.94	651	13.9
Infinitive	475	15.28	473	10.1
Imperative	279	8.78	609	13
Undefined modal	249	7.98	572	12.2
Modal cluster	1281	41.18	2305	49.2
Past	225	7.24	635	13.6
Present	480	15.4	868	18.6
Undefined present/past	21	0.68	67	1.4
Want, can	1102	35.46	806	17.2

To determine the significance of the differences between the three temporal categories in CDS and CS (plus the Undefined category in CS), several measures were taken. The proportions of occurrence of each temporal category were calculated for each day of recording in CDS and CS of each dyad, and cubic polynomial fits which capture their change over time were produced (Figure 1) revealing the same hierarchy of the temporal categories as depicted in Tables 8-11, where the Modal Cluster is more frequent than Present Tense, which is in turn more frequent than Past Tense in CDS and CS, and the Modal Cluster and Present Tense had a higher type and token frequency compared to the Undefined category in CS. The differences between Past Tense and the Undefined category in CS are less clear. The parents' use of the three temporal categories is relatively stable over time while the children's use shows changes, especially a decline in the use of the Undefined verbs and a slight increase in the use of Past Tense verbs.

INSERT FIGURE 1 ABOUT HERE

Kruskal Wallis (KW) tests revealed significant differences between the frequencies of the temporal categories in both types and tokens for the two dyads (Table 12). Post hoc Wilcoxon pairwise analyses of the proportions per temporal category were carried out to determine differences between pairs of categories, revealing that most of the differences were significant ($p<0.0001/0.01$), except Present vs Undefined and Past vs Undefined token frequencies in boy and girl CS ($z=-2.13$ $p<0.03$, $z= 1.77$ $p<0.0756$, $z=1.3$ $p<0.19$, $z=-0.62$ $p<0.05$, respectively), and Past vs Undefined token frequency in boy CS ($z=0.24923$ $p<0.8032$).

Table 12. KW test $\chi 2$ values ($p<0.0001$)

	Boy CDS		**Girl CDS**		**Boy CS**		**Girl CS**	
	type	**token**	**type**	**token**	**type**	**token**	**type**	**token**
$\chi 2$	113.4	117.4	109.1	122.8	118.7	100.6	112.4	91.5

Discussion

The current study presented a learning challenge in early Hebrew verb acquisition: How do children identify the root-and-pattern structure inherent to Hebrew verbs, given the pervasive opacity and inconsistency of verb forms in input to toddlers. The study provided a detailed account of the distribution of root types and temporal categories in Hebrew CDS and CS to show how verb specific morphological features act as distributional cues facilitating acquisition.

The role of type and token frequency is well known in language typology and grammatical learning. Natural languages are characterized by a high type frequency of regular forms (Lieven, 2010) and a high token frequency of irregular forms (Bybee, 2006). High type frequency is responsible for categorization while high token frequency is responsible for entrenchment of these forms. When dealing with early Hebrew verb acquisition, frequency alone does not provide a sufficient explanation for the way children break into this system, since verbs need not only to be learned as lexical items but, more importantly, to be analyzed into their crucial root and pattern components. However this non-linear composition is exactly what is masked in high-frequency, opaque verb forms.

Two main findings of the current study support our hypotheses regarding the way Hebrew acquiring children break into the verb system. First, our findings constitute new evidence that Hebrew verb CDS is fraught with defective (i.e., irregular) roots, which give rise to opaque verb forms - thus corroborating the premises of the current study. The data confirmed that defective root tokens prevailed over full root tokens in CDS, presenting toddlers with many more opaque verb forms obscuring the root and pattern combination that is critical for the acquisition of both verb inflection and derivation. On the other hand, type frequency of full roots was higher than that of defective roots, indicating that there are more full roots the child needs to learn as the highway to Hebrew verb morphology and lexical new-verb learning.

Second, our introduction pointed at a solution to this learning problem in the form of inflectional affixes in the Modal Cluster (Infinitive, Imperative, and Future Tense forms) as possible saliency-enhancing anchor points highlighting the internal structure of the stem. The fact that modal verbs serve as a central theme in early parent-child interaction (Aikhenvald, 2010; Stephany, 1983) supported our hypotheses. The prevalence of Modal Cluster verbs with salient boundaries was also confirmed in the current study.

Our findings confirmed the viability of this solution, as Modal Cluster verbs were found to be significantly more frequent in both CDS types and tokens, as compared to Present and Past Tense stems. Modal Cluster stems present the child with stable boundaries in the form of inflectional prefixes and suffixes highlighting their varying inner structure. Such boundaries may be regarded as a form of non-adjacent dependency, constituting a salient environment for verb learning (Gómez, 2002, 2008; Mintz, 2003, 2006; Sandoval & Gómez, 2013).

Children's output greatly resembled parental input distributions of roots, root types and temporal categories. They had a higher type frequency of full roots, a higher token frequency of defective roots, and a higher type and token frequency of Modal Cluster stems as compared to Present and Past Tense verbs, highlighting the role of CDS in shaping CS verb structure, and the possible contribution of CS to CDS verb content in the form of parental expansions and reformulations of children's utterances (Clark & De Marneffe, 2010).

These results advance our understanding of how Hebrew acquiring toddlers break into the verb system based on two structural characteristics of the verbs they hear in the input. First, the use of defective root verbs, mainly in the *Qal* verb pattern, resulting in verb lemmas such as *want, eat, see, come, bring* and *do* (Table 7). These verb meanings are not unique to Hebrew child language (Uziel Karl, 2003; Ninio, 2008; Clark & De Marneffe, 2010), however their prevalence in Hebrew CDS greatly contributes to verb opacity at a time when children's imminent task is to learn verb structure.

Second, the use of verbs with modal meaning testifying to the centrality of modality in early parent-child interaction is, again, not restricted to Hebrew (Aikhenvald, 2010; Stephany, 1983). Nonetheless, these general pragmatic properties of early verb use serve, again, as a helpful feature in figuring out how Hebrew verbs are structured. Thus, the solution to the proposed leaning problem in Hebrew verb acquisition actually lies in the combination of Hebrew-specific variegated defective stems with stable boundaries in the Modal Cluster coupled with universal propensities in the verb semantics and pragmatics of early parent child interaction.

Further research of the present data as well as data from older children and the input they receive is necessary on order to trace the future development of Hebrew verb acquisition. We predict that with time, Hebrew-learning children will gradually be exposed to changing verb distributions, as follows. Structurally, they will encounter a growing number of verbs with full roots, and temporality-wise, they will be exposed to more Present Tense forms expressing states and ongoing actions, as well as Past Tense forms carrying telic and then narrative meanings. These expectations are based on general learning trajectories of verb semantics in child language acquisition (Weist, 2009) as well as previous Hebrew studies (Berman, 1985 a, b). We expect a growing prevalence of more regular verb forms, based on full roots in the more transparent non-modal temporal stems (Ravid, 2012) in CDS targeting older children and in their own productions. This combination is regarded as the platform for learning the canonical patterns of Hebrew verb stems in each *binyan* and the consolidation of Hebrew verb architecture, eventually leading to command of non-linear root and pattern morphology across other lexical classes (Berman, 1987; Bolozky, 2007; Ravid, 2006; Schwarzwald, 2002).

The current study investigated the role of language-specific and universal features in CDS-CS relations, in line with similar research in other languages (Aksu-Koc, 1998; Xanthos et al., 2011). Specifically, it related Hebrew-specific and universal semantic-pragmatic properties of the verb in CDS as a solution to an initial learning challenge toddlers face in the process of Hebrew verb acquisition – how to extract the essential components of verb morphology from largely opaque verb forms. The current study not only identified this challenge for the first time, but also pointed the way for children to overcome it, based on the combination of three features separately documented in the literature: the prevalence of irregular over regular verb forms, the role of non-adjacent dependencies, and the dominance of modality semantics in CDS and CS (Bybee, 2006; Mintz, 2003, 2006; Stephany, 1983; Aikhenvald, 2010).

Acknowledgement: This work was supported by the Israel Science Foundation [grant number 285/13].

References

Aikhenvald, A. Y. (2010). *Imperatives and commands*. Oxford: Oxford University Press.

Aksu-Koç, A. (1998). The role of input vs. universal predispositions in the emergence of tense-aspect morphology: Evidence from Turkish. *First Language*, 18, 255-280

Albert, A., MacWhinney, B., Nir, B., & Wintner, S. (2012). The Hebrew CHILDES corpus: Transcription and morphological analysis. *Language Resources and Evaluation, 47*(4), 973-1005.

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 the Hebrew verb.* (PhD), Tel Aviv University / University of Antwerp, Tel Aviv / Antwerp.

Ashkenazi, O., Ravid, D., & Gillis, S. (Submitted). Input-output relations in the early acquisition of Hebrew verbs – morphological, semantic and pragmatic features.

Berman, R. A. (1975). The realization of syntactic process in the *binyan* system [in Hebrew]. *Hebrew Computational Linguistics, 9*, 25-39.

Berman, R. A. (1982). Verb pattern alternation: The interface of morphology, syntax and semantics in Hebrew child language. *Journal of Child Language, 9*, 169-191.

Berman, R. A. (1985). Acquisition of Hebrew. In D. I. Slobin (Ed.), *Crosslinguistic study of language acquisition* (pp. 255-371). Hillsdale: Lawrence Erlbaum.

Berman, R. A. (1985). *Acquisition of tense-aspect by Hebrew-speaking children. Final Report.* Binational Science Foundation (BSF) US-Israel.

Berman, R. A. (1986). A step-by-step model of language learning. In I. Levin (Ed.), *Stage and structure: Re-opening the debate* (pp. 191-219). Norwood: Ablex.

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

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

Berman, R. A. (1993). Developmental perspectives on transitivity: A confluence of clues. In Y. Levi (Ed.), *Other children, other languages: Issues in the theory of language acquisition* (pp. 189-240). Hillsdale: Lawrence Erlbaum.

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

Berman, R. A., & Armon-Lotem, S. (1996). How grammatical are early verbs? In C. Martinot (Ed.), *Actes du Colloque International sur l'Acquisition de la Syntaxe en Langue Maternelle et en Langue Étrangère* (pp. 17-60). Besançon: Presses Universitaires Franche-Comté.

Berman, R. A., Nayditz, R., & Ravid, D. (2011). Linguistic diagnostics of written texts in two school-age populations. *Written Language & Literacy, 14*, 161-187.

Berman, R. A., & Slobin, D. I. (1994). *Relating events in narrative: A crosslinguistic developmental study*. Lawrence Erlbaum.

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

Bittner, D., Dressler, W. U., & Kilani-Schoch, M. (Eds.). (2003). *Development of verb inflection in first language acquisition – A cross linguistic perspective*. New York: De Gruyter Mouton.

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

Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language, 82*, 711-733.

Bybee, J. (2007). *Frequency of use and the organization of language*. Oxford: Oxford University Press.

Clark, E. V., & de Marneffe, M. C. (2012). Constructing verb paradigms in French: Adult construals and emerging grammatical constructs. *Morphology, 22,* 89-120.

Davies, M. (2009). The 385+ million word corpus of contemporary American English (1990-present). *International Journal of Corpus Linguistics, 14*, 159-190.

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

Ellis, N. (2012). Emergentism. In C. Chapelle, M. Gullberg, & J. Williams (Eds.), *The encyclopedia of applied linguistics: Cognitive approaches*. Hoboken: Wiley-Blackwell.

Gentner, D. (1982). Why nouns are learned before verbs. In S. Kuczaj (Ed.), *Language development Vol. 2: Language thought and culture* (pp. 301-334). Hillsdale: Lawrence Erlbaum.

Golinkoff, R. M., & Hirsh-Pasek, K. (2006). Introduction: Progress on the verb learning front. In R. M. Golinkoff & K. Hirsh-Pasek (Eds.), *Action meets words: How children learn verbs* (pp. 3-28). Oxford: Oxford University Press.

Golinkoff, R. M., & Hirsh-Pasek, K. (2008). How toddlers begin to learn verbs. *Trends in Cognitive Science, 12*, 397-403.

Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science, 13,* 431-436.

Gómez, R. L. (2008). Statistical learning in infant language development. In M. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 601-616). Oxford: Oxford University Press.

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

Herzberg, O. (2010). *Early morphosyntactic and lexical developmentin young hearing impaired Hebrew-speaking children with cochlear implants compared to normally hearing peers.*

(PhD), Tel Aviv University / University of Antwerp, Tel Aviv /Antwerp.

Kaplan, D. (1983). *Developmental trends in the grammar of 2-3 year old children [in Hebrew]*. (MA), Tel Aviv University, Tel Aviv.

Kibrik, A. (2012). Toward a typology of verbal lexical systems: A case study in Northern Athabaskan. *Linguistics*, *50*, 495-532.

Koptjevskaja-Tamm, M. (2012). New directions in lexical typology. *Linguistics*, *50*, 373-394.

Kuhl, P. (2000). A new view of language acquisition. *PNAS, 97*, 1850-1857.

Kuhl, P. (2004). Early language acquisition – cracking the speech code. *Nature Reviews Neurosciences, 51*, 831-843.

Lieven, E., Behrens, H., Speares, J., & Tomasello, M. (2003). Early syntactic creativity: A usage based approach. *Journal of Child Language, 30*, 333-370.

Lieven, E. (2008). Learning the English auxilliary: A usage based approach. In H. Behrens (Ed.), *Corpora in language acquisition research – History, methods, perspectives* (pp. 61-98). Amsterdam: Benjamins.

Lieven, E. (2010). Input and first language acquisition: Evaluating the role of frequency. *Lingua*, *120*, 2546-2556.

Lustigman, L. (2012). Non-finiteness in early Hebrew verbs. *BAALL, 4*, 213-231.

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

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*. Hillsdale: Lawrence Erlbaum.

Maital, S., Dromi, E., Sagi, A., & Bornstein, M. (2000). The Hebrew Communicative Development Inventory: Language specific properties and cross linguistic generalizations. *Journal of Child*

Language, 27, 43-67.

Mintz, T. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition, 90*, 91-117.

Mintz, T. (2006). Finding the verbs: Distributional cues to categories available to young learners. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets words: How children learn verbs* (pp. 31-63). Oxford: Oxford University Press.

Nazzi, T., & Houston, D. (2006). Finding verb forms within the continuous speech stream. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets words: How children learn verbs* (pp. 64-87). Oxford: Oxford University Press.

Ninio, A. (2008). Language and the learning curve: A new theory of syntactic development. *Journal of Child Language, 35*, 922-929.

Romberg, A., & Saffran, J. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*, 906-914.

Saffran, J. (2003). Statistical language learning: Mechanism and constraints. *Current Directions in Psychological Science, 12*, 110-113.

Sandoval, M., & Gómez, R. L. (2013). The development of nonadjacent dependency learning in natural and artificial languages. *Wiley Interdisciplinary Reviews: Cognitive Science, 4*, 511-522.

Ravid, D. (1995). *Language change in child and adult Hebrew: A psycholinguistic perspective*. Oxford: Oxford University Press.

Ravid, D. (1997). Early morphological development a duo: Pre- to proto morphology in Hebrew-speaking twins. *Papers and Studies in Contrastive Linguistics, 33*, 79-102.

Ravid, D. (2003). A developmental perspective on root perception in Hebrew and Palestinian Arabic. In J. Shimron (Ed.), *Language processing and acquisition of Semitic, root-based, morphology* (pp. 293-319). Amsterdam: Benjamins.

Ravid, D. (2006). Word-level morphology: A psycholinguistic perspective on linear formation in Hebrew nominals. *Morphology, 16*, 127-148.

Ravid, D., Ashkenazi, O., Levie, R., Ben Zadok, G., Grunwald, T., Bratslavsky, R., & Gillis, S. (In press). Foundations of the root category: Analyses of linguistic input to Hebrew-speaking children. In R. A. Berman (Ed.), *Acquisition and development of Hebrew: From infancy to adolescence*. Amsterdam: Benjamins.

Ravid, D., & Schiff, R. (2015). It's all about gender: Hebrew speakers' processing of plural agreement morphology. *Morphology, 25*, 327-343.

Rose, T., Stevenson, M., & Whitehead, M. (2002). The Reuters Corpus Volume 1. *Language Resources and Evaluation, 2,* 817-832.

Schachter, P. (1985). Parts of speech systems. In T. Shopen (Ed.), *Language typology and syntactic description: Clause structure* (pp. 3-61). Cambridge: Cambridge University Press.

Stephany, U. (1983). *The development of modality in language acquisition*. Köln: Arbeitspapier / Institut für Sprachwissenschaft.

Stumper, B., Bannard, C., Lieven, E., & Tomasello, M. (2011). "Frequent frames" in child directed speech: A limited cue to grammatical categories. *Cognitive Science, 35,* 1190-1205.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and syntactic description – Grammatical categories and the lexicon* (pp. 57-138). Cambridge: Cambridge University Press.

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

Tomasello, M. (2009). The usage based theory of language acquisition. In E. Bavin (Ed.), *The Cambridge handbook of child language* (pp. 69-88). Cambridge: Cambridge University Press.

Tomasello, M. (2006). Acquiring linguistic constructions. In R. Siegler & D. Kuhn (Eds.), *Handbook of child psychology* (pp. 256-293). New York: Wiley.

Tomasello, M. (2004). Learning through others. *Dedalus, 133,* 51-58.

Uziel Karl, S. (2003). Motion verbs in the developing lexicon of Hebrew speaking children. In E. V. Clark (Ed.), *Proceedings of the 31st Stanford Child Language Research Forum* (pp. 31-39). Stanford: CSLI Publications.

Veneziano, E., & Parisse, C. (2010). The acquisition of early verbs in French: Assessing the role of conversation and of child-directed speech. *First Language*, 30, 287-311.

Weist, R. (2009). Children think and talk about time and space. In P. Łobacz, P. Nowak, & W. Zabrocki (Eds.), *Language, science and culture*. Poznań: Wydawnictwo Naukowe UAM.

Xanthos, A., Laaha, S., Gillis, S., Stephany, U., Aksu-Koç, A., Christofidou, A... & Korecky-Kröll, K. (2011). On the role of morphological richness in the early development of noun and verb inflection. *First Language*, 20, 1-38.

Yurovsky, D., Yu, C., & Smith, L. B. (2012). Statistical speech segmentation and word learning in parallel: Scaffolding from child-directed speech. *Frontiers in Psychology, 3*(374).

Endnotes

[1] There are in fact seven *binyan* patterns, but the strictly passive *Pu'al* and *Huf'al* are irrelevant to the age range in the current study.

[2] 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).

[3]Present tense agrees only with number and gender, not with person

[4] Most future forms serve also for the imperative function (Bolozky, 1979)

[5]Infinitives have a modal or irrealis meaning (Hyams, 1998), imperatives are modal in nature as they convey requests or commands, closely related to deontic modality (Palmer, 1986) and future tense verbs are usually inherently uncertain to actuality, making them mostly modal in nature (Chung & Timberlake 1985). Furthermore, "*futurity is never purely a temporal concept … what is conventionally used as a future tense … is rarely if ever used solely for making statements or predictions ... about the future. It is also used … in utterances involving supposition, inference, wish, intention and desire*" Lyons (1977).

[6] Verbs such as *rotse* / *rotsa* 'want, Present, SG, MS / FM' provided a felicitous environment for the disambiguation of truncated verbs as probable infinitival forms. However, such truncations were marked as Infinitive only if closely resembling the infinitival form. Consider, for example, *ani otse esaxek* 'I want play (truncated)' where *esaxek* closely resembles *lesaxek* 'play, Inf' in the appropriate syntactic environment. Each occurrence of such truncated verbs was checked separately by the first two authors, with disagreements resolved by discussion.

---

[7] The option of *tagid* standing for 3rd, MS/FM,SG' 'she will say' does not seem viable here.

[8] As the token frequency of Present Tense verbs was biased by an extremely high occurrence of modal ‘be able’ and cognitive 'want', especially in CS, CS token data is presented without these verbs.

References

Aikhenvald, A.Y. (2010)**.** *Imperatives and commands***.** Oxford University Press**.**

Albert, A., MacWhinney, B., Nir. B. & Wintner, S. (2012)**.** The Hebrew CHILDES corpus transcription and morphological analysis. *Language Resources and Evaluation*, 1-33.

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 the Hebrew verb*. An unpublished doctoral dissertation. Tel Aviv University.

Ashkenazi, O., Ravid, D., Gillis, S. (submitted**).** Input-output relations in the early acquisition of Hebrew verbs – morphological, semantic and pragmatic features.

Berman, R. A. *(1975).* The realization of syntactic process in the *binyan* system. *Hebrew Computational Linguistics 9, 25-39.* [Hebrew]

Berman, R.A. (1982). Verb pattern alternation: the interface of morphology, syntax and semantics in Hebrew child language. *Journal of Child Language*, 9, 169-191.

Berman, R. A (1985). Acquisition of Hebrew.I n D. I. Slobin, (Ed). *Crosslinguistic Study of Language Acquisition,* Volume I (pp. 255-371), Hillsdale, NJ: Lawrence Erlbaum

Berman, R. (1985) Acquisition of tense-aspect by Hebrew-speaking children. *Final Report, Binational Science Foundation (BSF*) US-Israel.

Berman, R.A. (1986). A step by step model of language learning. In I. Levin (Ed.) *Stage and Structure: re-opening the Debat*e (pp. 191-219). Norwood, N.J: Albex.

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

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

Berman, R.A. (1993). Developmental perspectives on transitivity: A confluence of clues. In Y. Levi, (Ed) *Other Children, Other Languages: Issues in the Theory of Language Acquisition* (pp. 189-240). Hillsdale, NJ: Earlbaum.

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

Berman, R.A., Armon Lotem S. (1996). How grammatical are early verbs? In C. Martinot (Ed.) *Annales Littéraires de l'Université de Franche-Comté: Actes du Colloque International sur l'Acquisition de la syntaxe* (pp.17-60).

Berman, R.A, Nayditz, R. & Ravid, D. (2011). Linguistic diagnostics of written texts in two school-age populations. *Written Language & Literacy* 14, 161-187.

Berman, R.A & Slobin, D. (1994). Relating Events in Narrative: A Crosslinguistic Developmental Study. Hillsdale, NJ: Lawrence Erlbaum.

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

Bittner, D., Dressler, W.U., Kilani-Schoch, M. (2003). *Development of verb inflection in first language acquisition – a cross linguistic perspective*. Mouton de Gruyter, new York.

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

Bybee, J. (2006). From usage to grammar: the mind's response to repetition. *Language* 82, 711-733.

Bybee, J. (2007). *Frequency of use and the organization of language*. Oxford: Oxford University Press.

Clark, E. V., de Marneffe, M.C (2012) Constructing verb paradigms in French: adult construals and emerging grammatical constructs. *Morphology*, 22, 89-120.

Davies, M. (2009). The 385+ Million Word Corpus of Contemporary American English (1990-present*). International Journal of Corpus Linguistics* 14, 159-90

Dromi, E. (1987). *Early lexical development.* Cambridge: Cambridge university press

Ellis, N.C (2012). Emergentism. In C. A. Chapelle, M. Gullberg, J.Williams (Eds.). *The Encyclopedia of Applied Linguistics: Cognitive Approaches. Wiley-Blackwell.*

Gentner,D. (1982). Why nouns are learned before verbs. In S. Kuczaj II (Ed.), *Language development Vol. 2: Language thought and culture* (pp. 301-334). Hillsdale, N.J.: Lawrence Erlbaum.

Golinkoff, R.M., Hirsh-Pasek, K. (2006). Action abets accomplishment: Progress on the Verb learning Front. In R.M. Golinkoff & K. Hirsh-Pasek (Eds.) *When*

Action Meets Words: How young children learn Verbs (pp. 1-48). Oxford University Press.

Golinkoff, R., & Hirsh-Pasek, K. (2008). How toddlers learn verbs. *Trends in Cognitive Science*, 12(10), 397-403.

Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, 13, 431-436.

Gómez. R.L. (2008). Statistical learning in infant language development. In M.G. Gaskell (Ed). *The Oxford Handbook of Psycholinguistics* (pp. 601-616.). Oxford University Press.

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

Herzberg, O. (2010). *Early morphosyntactic and lexical developmentin young hearing impaired Hebrew-speaking children with cochlear implants compared to normally hearing peers.* Unpublished PhD. Dissertation, Tel Aviv University.

Kaplan, D. 1983. *Developmental trends in the grammar of 2-3 year old children*. Unpublished MA Thesis. Tel Aviv University. [Hebrew].

Kibrik, A.A. 2012. Toward a typology of verbal lexical systems: A case study in Northern Athabaskan. *Linguistics* 50(3), 495 – 532.

Koptjevskaja-Tamm, M. 2012. New directions in lexical typology. *Linguistics*, 50(3), 373-39.

Kuhl, P. (2000). *A new view of language acquisition*. PNAS, 97 (22), 1850-1857.

Kuhl, P. (2004). Early language acquisition – cracking the speech code. *Nature Reviews Neurosciences*, 51, 831-843.

Lieven, E., Beherens, H., Speares, J. & Tomasello, M. (2003). Early syntactic creativity: A usage based approach. *Journal of Child Languag*e, 30, 333-370.

Lieven, E. (2008). Learning the English auxillary: A usage based approach. In Beherns, H. (Ed.) *Corpora in language acquisition research – History, methods, perspectives (pp. 61-98). John Benjamin Publications.*

Lieven, E., (2010). Input and first language acquisition: evaluating the role of frequency. Lingua. 120, 2546-2556.

Lustigman, L. (2012). Non-finiteness in early Hebrew verbs. *BAALL* 4:213-231.

Lustigman, L. (2013). Developing structural specificification: Productivity in early Hebrew verb usage. *First Language*, 33(1), 47-67.

MacWhinney, B. (2000). The CHILDES project: Tools for analyzing talk. Hillsdale NJ: Lawrence Erlbaum Associates.

Maital S., Dromi E., Sagi A, Bornstein M.H (2000). The Hebrew Communicative Development Inventory: Language specific properties and cross linguistic generalizations. Journal of Child Language, 27, 43-67.

Mintz, T. (2003). Frequent frames as a cue for grammatical categories in child cirected speech. *Cognition*, 90, 91-117.

Mintz T.H. (2006). Finding the verbs: distributional cues to categories available to young learners. In Hirsh Pasek. K. & Golinkoff, R.M (Eds). *Action Meets Words, How Children Learn Verbs* (pp. 31-63). Oxford University Press.

Nazzi, T., Houston, D. (2006). Finding verb forms within the continuous speech stream. In Hirsh Pasek K. & Michnick Golinkoff R. M (Eds). *Action meets words: How children learn verbs*, (pp. 64-87). Oxford University Press.

Ninio, A. (2008). Language and the learning curve: A new theory of syntactic development. *Journal of Child Language*, 35, 922-929.

Romberg A.R., Saffran J.R. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(6), 906-914.

Saffran, J. R (2003). Statistical language learning: mechanism and constraints. *Current Directions in Psychological Science*,12 (4), 110-113.

Sandoval, M., Gómez, R.L. (2013). The development of non adjacent dependency learning in natural and artificial languages. *WIREs Cognitive Science*, 4, 511-522.

Ravid, D. (1995). *Language Change in Child and Adult Hebrew: A Psycholinguistic perspective*. New York: Oxford University Press.

Ravid, D. (1997). Early morphological development a duo: pre- to proto morphology in Hebrew speaking twins, *Papers in the Study of Contrastive Linguistics*, 33, 79-102.

Ravid, D. (2003). A developmental perspective on root perception in Hebrew and Palestinian Arabic. In Shimron J. (Ed) *Language Processing and Acquisition of Semitic Root-based Morphology* (pp. 293-320). John Benjamin Pub Ravid, D. (2010). The emergence of the Hebrew verb category. *Literacy and Language*, 3. [Hebrew]

Ravid, D. (2006). Word-level morphology: A psycholinguistic perspective on linear formation in Hebrew nominals. *Morphology*, 16, 127-148.

Ravid, D. (2012). *Input-output patterns in the acquisition of Hebrew root usage: A corpus-based psycholinguistic study.* An ISF research proposal, Research Grant Application no. 285/13.

Ravid, D., Ashkenazi, O., Levie, R., Ben Zadok, G., Grunwald,T., Bratslavsky R.and Gillis, S. (In press). Foundations of the root category: analyses of linguistic input to Hebrew-speaking children. *In R. Berman (Ed.) Acquisition and Development of Hebrew: From Infancy to Adolescence. Amsterdam: John Benjamins.*

Ravid, D. & Schiff, R. (2015). It's all about gender: Hebrew speakers' processing of plural agreement morphology. *Morphology*, 25, 327–343.

Rose, T., Stevenson, M., & Whitehead, M. (2002). The Reuters Corpus Volume 1. *LREC* ,2, 827-832.

Schachter, P. (1985). Parts of speech systems. In: T. Shopen (Ed.), *Language typology andsyntactic description: Clause structure* (pp. 3-61). Cambridge: Cambridge University Press.

Schwarzwald, O.R. 2002. *Modern Hebrew morphology*. Tel Aviv: Open University [Hebrew].

Schwarzwald R. O. (1998). Inflection and derivation in Hebrew linear word formation. *Folia Linguistica*, 1998, 32, 265-287.

Stephany, U. (1983). *The development of modality in language acquisition*. A working paper, Linguistic institute, Koln University.

Stumper, B., Bannard, C., Lieven, E., Tomasello, M. (2011**).** "Frequent frames" in child directed speech: a limited cue to grammatical categories. *Cognitive Science*, 35, 1190-1205.

Talmy, L. (1985). Lexicalization patterns: semantic structure in lexical forms. In *T. Shopen (Ed.) Language Typology and Syntactic Description – Grammatical Categories and the Lexicon (pp. 57-138). Cambridge University Press.*

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

Tomasello, M. (2009). The usage based theory of language acquisition. In E. L. Bevin(Ed), *The Cambridge Handbook of Child Language* (pp. 69-88), Cambridge University Press.

Tomasello, M. (2006). Acquiring linguistic constructions. In R., Siegler & D. Kuhn (Eds) *Handbook of Child Psychology* (pp. 256-293). New York: Wiley.Kibrik, 2012

Tomasello, M. (2004). Learning through others. *Dedalus*, 133, 51-58.

Uziel Karl, S. (2003). Motion verbs in the developing lexicon of Hebrew speaking children. In: U.,Ornan, R., Ben Shahar, and G. Turi (Eds.) *Hebrew as a living language: Research on language in social-cultural context.* Following the Oranim conference. Hakibuts Hameuhad, Tel Aviv University [Hebrew].

Weist, R. M. (2009). Children think and talk about time and space. In P. Lobacz, P. Nowak, & W. Zabrocki (Eds.) *Language, science and culture*. Poznań: Wydawnictwo Naukowe UAM.

Yurovsky, D., Yu, C., Smith, L. B (2012). Statistical speech segmentation and word learning in parallel: scaffolding from child-directed speech. Frontiers in psychology, 3, 1-9.