Hebrew adjective lexicons in developmental perspective

Subjective register and morphology

Dorit Ravid, Amalia Bar-On, Ronit Levie, and Odelia Douani
Tel Aviv University

Objective frequency does not always provide reliable information about lexical distributions across individuals' development. We propose the subjective ranking by experts of lexical items' register in the sense of 'levels of linguistic usage', which has been independently linked to AoA, as an alternative. This proposal was tested in Hebrew, a language showing marked distinctions between the everyday colloquial style and more formal, historically-related types of expression. A list of over 3,500 Hebrew adjectives in 19 morphological categories was compiled from dictionary sources. All adjectives on the list were ranked on a 1–5 linguistic register scale by 329 language expert judges. A Model Based Latent Class Analysis yielded five high-agreement groups of adjectives with mean register scores from 1.44 to 4.51, taken to represent five developmentally consecutive adjective lexicons. Semantic and morphological analyses indicated a rise in the abstractness and specificity of adjectives in the five lexicons, with concurrent changes in their morphological makeup. Two morphological categories emerged as the major components of the Modern Hebrew adjective lexicon: Resultative patterns, expressing states, and *i*-suffixed denominals, expressing nominal attributes. The study showed that subjective register classification may constitute a yardstick in development, with implications for other languages where register judgements can apply.

Keywords: adjectives, morphology, AoA, subjective ranking, register, Hebrew

Most psycholinguists would agree that gaining information about lexical knowledge in learners is critical for the construction of theories of language acquisition (Douglas, 2003; Kennedy, 2014). *Frequency* has been a cornerstone of psycholinguistic research on the lexicon, as well as other domains of language learning, usage and change (Diessel, 2007; Gries & Divjak, 2012; Jurafsky, 2003). However,

The Mental Lexicon 11:3 (2016), **401–428.** DOI 10.1075/ml.11.3.04rav
ISSN 1871–1340 / E-ISSN 1871–1375

in seeking reliable information about lexical distributions of words across individuals' development, corpora frequency has been found wanting (Balota, Pilotti, & Cortese, 2001; Brysbaert & Ghyselinck, 2006). We propose an alternative to objective frequency counts in corpora – subjective ranking by experts of lexical items' *register* in the sense of 'levels of linguistic usage', which has been independently linked to AoA (Ravid & Berman, 2009; Ravid & Vered, in press).

Adjectives

Adjectives are of particular interest in vocabulary development as the smallest (often absent) and most diverse lexical category in many languages (Dixon & Aikhenvald, 2014; Schachter & Shopen, 2007). As relational terms, adjectives show up later on in child speech than nouns and verbs (Casseli et al., 1995; Salerni et al., 2007), and they constitute a low-frequency class compared to other content words in children's early lexicons (Tribushinina et al., 2014; Sandhofer, Smith, & Luo, 2000). A full array of adjectival categories is far from present even in 6-year-olds (Blackwell, 2005; Blodgett & Cooper, 1988), suggesting it coincides with the consolidation of a literate lexicon and its cognitive correlates (Dockrell & Messer, 2004; Ravid & Levie, 2010). The size and makeup of the adjective category can thus be taken as a yardstick for language development and proficiency. A first step towards this goal would be to amass a corpus of a language's adjectives and classify it according to a well-motivated criterion that would reflect developmental goals. Given the morphological richness of Hebrew, adjective morphology would be a first choice for such a developmentally-oriented classification.

The morphology of Hebrew adjectives

Modern Hebrew adjectives are particularly diverse morphologically. They make use of all three morphological structures in the language: Non-linear (root-and-pattern) forms, linear (stem-and-suffix) forms, and reduplicative forms.

Non-linear adjective morphology

Many adjective forms are based on the typically Semitic discontinuous root and pattern structure – both verbal or nominal (Ravid, 1990, 2003; Ravid & Schiff, 2006), e.g., *menumas* 'polite' (root *n-m-s,* pattern[1] *meCuCaC*). *Verbal patterns.* The participial (so-called *beynoni* 'middle') patterns of the seven *binyan* verb con-

1. Patterns are represented as templates with C's denoting root radicals (Ravid, 1995).

jugations constitute one of the central word-formation mechanisms in Hebrew (Berman, 1978; Ravid, 2012). In addition to their inflectional role as present-tense temporal stems (Ravid et al., 2016), *beynoni* patterns productively derive adjectives or nouns (Berman, 1993). For example, *madhim* 'amazing' (root *d-h-m,* pattern *maCCiC*) serves not only as the present-tense stem in the temporal paradigm of the *hif'il* verb 'amaze', but also as an adjective in its own right.

Verb-related present-tense patterns fall into two distinct groups. One consists of the three resultative passive participle patterns *CaCuC* (*saduk* 'cracked', the resultative counterpart of the *qal* and *nif'al* conjugations), *meCuCaC* (*mekulkal* 'spoiled', the present-tense form of the *pu'al* conjugation), and *muCCaC* (*mus-max* 'qualified', the present-tense form of the *huf'al* conjugation) (Berman, 1994; Ravid, 2004). A second group consists of the *beynoni* patterns of the five non-passive *binyan* conjugations – *CoCeC*[2] (*qal*, e.g., *bolet* 'prominent'), *niCCaC*[3] (*nif'al*, e.g., *ne'eman* 'loyal'), ma*CCiC* (*hif'il,* e.g., *maksim* 'charming'), *meCaCeC* (*pi'el*, e.g., *medake* 'depressing'), and *mitCaCeC* (*hitpa'el*, e.g., *mictayen* 'outstanding').

Nominal patterns

Hebrew Adjectives also productively share patterns with nouns. These are mainly *CaCCan* (e.g., *akshan* 'stubborn', cf. noun *safran* 'librarian'), and *CaCiC* (e.g., *ta'im* 'tasty', cf. *pakid* 'clerk'), as well as less productive nominal patterns such as *CaCaC* (adjective *yashar* 'honest', noun *naxhash* 'snake'), the traditional category of handicaps *CiCeC* (e.g., *iver* 'blind'), and *CaCoC* (*gadol* 'big'), including color adjectives (e.g., *yarok* 'green').

Linear adjective morphology

Denominal adjective derivation (e.g., *kav-i* 'line-ar'), estimated to include over 1/3 of the adjectives in Hebrew (Ravid & Shlesinger, 1987), is part of Hebrew nominal word-level morphology expanding word stems by linear suffixation (Ravid, 2006). Modern denominal adjectives originate in a small Classical Hebrew class of *i*-suffixed nouns denoting ethnic origin, e.g., *yevusi* 'of Yevus'. These evolved in Medieval Hebrew and even more so in Modern Hebrew into the large, full-fledged class of adjectives derived by the suffixation of *-i* to a noun base (Berman, 1978).

---

2. As the semantically basic, most prevalent and most structurally diverse *binyan* conjugation, *Qal* has several present-tense allomorphs in addition to the default *CoCeC*, the most prominent of which is *CaCeC,* e.g., *ayef* 'tired'.

3. Although *nif'al* is considered a passive *binyan* (e.g., *nixtav* 'written'), much evidence points to its having a middle, non-agentive character such as *hitpa'el* (Berman, 1993; Ravid & Vered, in press).

Reduplication
Reduplication involves the repetition of part of the stem or its entirety, e.g., Agta *ulu / ululu* 'head/s' (Spencer, 1991). In Hebrew it works from left to right, e.g., adjective *shaxor / shxarxar* 'black / darkish'. Compared with the two widespread non-linear and linear systems that serve Hebrew inflection and derivation (Ravid, 2003, 2006), reduplication constitutes a third and rather minor structural device, generally restricted to diminutive expression in nouns and adjectives[4] (Hora et al., 2006).

Adjectives in complex structures
Adjectives moreover participate in two complex structures – taking quantifying prefixes of Aramaic origin, as in *tlat-xodshi* 'thrice-monthly',[5] and adjective-headed compounds, as in *kcar ruáx* 'short (of) spirit = impatient' (Ravid & Levie, 2010).

Morphology as a developmental criterion

Studies have identified a connection between morphological adjective categories and the course of their acquisition in Hebrew. Accordingly, Ravid and Levie (2010) constructed a morpho-semantic adjective scale that took into account category semantics, structural complexity coupled with adjective distributions derived from naturalistic data and experimental elicitation. The lowest rank on this 1–4 scale consisted of a core list of mainly short, morphologically simplex Classical adjectives bearing primary adjectival meaning (*tov* 'good', *xam* 'hot'), prevalent in the speech of toddlers and their caregivers (Berman, 1985; Ravid & Nir, 2000; Tribushinina et al., 2014). Rank 2 on the scale constituted resultative passive adjectives such as *metukan* 'fixed', the earliest productive adjective category in preschool Hebrew (Berman, 1993; Ravid, 1995). Rank 3 consisted of adjectives with non-linear form, including non-passive *beynoni* participles (e.g., *macxik* 'funny'), nominal patterns (e.g., *shakran* 'liar'), and reduplicative adjectives (e.g., *CCaCCaC* (e.g., *shmanman* 'fatso'), occurring in the usage of school-going children (Ravid, Levie, & Avivi Ben Zvi, 2003). Rank 4 was dedicated to linear *i*-suffixed denominal adjectives such as *tipusi* 'typical', a later-developing, school-oriented category (Levin, Ravid, & Rapaport, 2001; Ravid & Zilberbuch,

4. Note however that this is not the full story, as reduplication processes can apply to derived roots (e.g., *k-d-r-r* 'dribble' from *kadur* 'ball', or *?-v-r-r* 'bring air in' from *avir* 'air').

5. We excluded prefixed adjectives whose heads can be used as independent adjectives without the quantifying prefix, e.g., *cdadi* 'side'.

2003), requiring familiarity with a large, multifaceted nominal lexicon typical of literary prose, informative and expository texts (Ravid, 2004).

Ravid and Levie's (2010) analysis of adjectives in a corpus of spoken and written Hebrew texts generally confirmed the 1–4 scalar order. Across the school years there was a steady age-related increase in adjective ranking on the scale, in number of adjective rank levels, adjective category size and semantic abstractness and sophistication of specific adjectives, especially between adolescence and adulthood, and mostly in written expositions. Denominal adjectives (rank 4) were clearly found to associate with older age, literacy and advanced text production skills. However, this classification also proved to be too coarse to detect reduplicative adjectives, also missing complex adjectives, which were included in the syntactic rather than morphological analyses. Most important, ranks 1–3 contained adjectives that could be associated with different ages and language skills, e.g., everyday *shavur* 'broken' and lexically specific *natuy* 'slanted' in rank 2. Moreover, distinct groups with different semantic values shared the same patterns, e.g., *CaCiC* with a basic early adjective such as *ta'im* 'tasty', but also a productive class of potential adjectives such as *axil* 'edible', atypical of young children's language. Thus, Hebrew morphological categories cannot serve as the sole markers of adjectival AoA.

To rank Hebrew adjectives in a way that would reliably reflect language richness or proficiency, it was necessary to create an independent measure that would take into account individual adjectives and cluster them into meaningful groups, while relating this measure to adjective morphological categories so as to determine the viability of the previous classifications.

Familiarity as index

Word frequency, associated with an individual's developing lexicon, would probably be the first choice of developmental researchers in classifying adjectives by AoA, as words that occur more frequently in the language also tend to be learned earlier on (Brysbaert & Cortese, 2011; Juhasz, 2005; Kuperman & van Dyke, 2013). Moreover, earlier lexical learning implies more exposure to the word in diverse communicative settings, and hence more familiarity with its specific shades of meaning, its morphological modulations, and its syntactic and pragmatic co-occurrence neighbors. One prominent way of gaining information about word frequency in the language is through *objective frequency* of occurrence in databases. However, objective frequency might not reflect the true distributions in the language because of a bias towards written over spoken sources (Balota et al., 2001) or towards a discourse genre such as sports, the news or the sciences (Biber & Conrad, 2009; Carroll, 1970). Important to the current analysis, objective

frequency is most often obtained from adult corpora, and therefore cannot reflect the linguistic experience of children and adolescents (Brysbaert & Ghyselinck, 2006). While the CHILDES databases provide excellent information on frequencies in child-directed and child speech, most of them concern young children in conversation with caregivers (MacWhinney, 2000). Thus, obtaining information from objective frequencies about the long developmental course over later childhood and adolescence of a category of words such as adjectives is not really an option.

Hebrew adjectives in texts

Attempts to construct a reliable adjective corpus in Hebrew have met with difficulties similar to those pointed above. For example, the Wintner Corpus (Itai & Wintner, 2008) is based on the 1990–1991 issues of the Israeli daily newspaper *Ha-Aretz,* and thus biased towards words frequent in the dated Hebrew lexicons of highly literate and sophisticated adult readers. Unsurprisingly, it does not contain adjectives such as *viral* or *autistic,* while *sovyeti* 'Soviet' occurs with high frequency (Benelli & Grinboim, 2015).

Obtaining an adjective list from Hebrew textual sources poses two additional, language-specific challenges. First, typological and historical factors have rendered the Hebrew adjective category extremely fuzzy. Classical Hebrew, lacking a dedicated morphological class of adjectives, used present tense verb participles and ethnic nominals to denote states, attributes and properties of nouns (Gai, 1995). As a result, current Hebrew adjectives not only share structural properties with both nouns and verbs, but in fact many adjectives are ambiguous across two or three content-word classes, defined only by syntactic factors in discourse, e.g., *metapes,* which might be interpreted as noun 'climber', present-tense verb 'climb/s', or adjective 'climbing'. Thus, a form such as *mag'il* might be the present-tense form in the verb paradigm where past and future tense would respectively mean 'caused / will cause disgust', or an adjective meaning 'disgusting'. Distinguishing adjectives from nouns and verbs makes the reliable identification of adjectives in corpora problematic. In fact, the Wintner Corpus contains about 1,000 words labeled as adjectives that do not appear in the Avneyon (2007) dictionary used in the current endeavor (Benelli & Grinboim, 2015). Such discrepancies would render a corpus-derived Hebrew adjective list rather unreliable.

A second Hebrew-specific problem is homography, which runs rampant in the default vowel-poor Hebrew orthography, so that a string such as לבנוני might be read as either *levanóni* 'Lebanese' or *lavnuni* 'whitish' (Bar-On & Ravid, 2011). The fact that several function words such as the definite article *ha* or the

subordinating marker *she* are attached to the following written word increases homography tenfold (Ravid, 2012). Disambiguating homographic strings requires detailed manual analyses taking into account the semantic and morpho-syntactic environment of a large proportion of the words under consideration (Bar-On, Dattner, & Ravid, submitted). Homography too contributes to the unreliability of objective frequency counts in written Hebrew corpora.

Subjective judgments of accrued experience in hearing, reading or using certain words constitute an alternative way of obtaining information about word frequency. True, subjective familiarity and frequency rating too may be contaminated by measures like the word's abstractness and orthographic similarity to other words (Thompson & Desrochers, 2009), inflectional and derivational entropies (Baayen, Feldman, & Schreuder, 2006), as well as contextual factors in the elicititation, such as clarity of meaning (Balota et al., 2001). Nonetheless, taking into account participants' age, educational backgrounds, literacy skills, clinical conditions and other variables, several studies have shown a strong relation between subjective frequency ratings and objective frequency estimates from corpora (Auer, Burnstein, & Tucker, 2000; Balota et al., 2004). In fact, as pointed out in Kuperman and Van Dyke (2013), subjective assessments of word frequency have been found to be superior over objective frequency gained from corpora.

In addition, unlike objective frequency, subjective frequency ratings can be elicited with language and literacy development in mind. The last years have seen the increasing use of language experts' subjective assessment of word familiarity as a viable alternative to objective frequency rating (Thompson & Desrochers, 2009). Importantly, the ability to make valid assessments is related to the linguistic and reading experience of raters (Auer, Burnstein, & Tucker, 2000; Kuperman & Van Dyke, 2013). Thus, more proficient language users make better judgment regarding word frequency, as they are familiar with more, and rarer words (Alderson, 2007). In the current study, we adopted linguistic *register* as a criterion for the subjective ranking of adjectives by language experts.

Register

Ferguson (1994: 16) defines register as "the linguistic differences that correlate with different occasions of use". This definition was extended to 'levels of linguistic usage' (Ravid & Berman, 2009) as developmentally, gaining command of linguistic register involves learning the range of expressive options available in the target language, and being able to map relevant linguistic forms in accordance with communicative context. Several studies have found a relationship between register sensitivity and language development (Andersen, 1992; Clark, 2003).

Ravid and Berman (2009) demonstrated that register can provide a yardstick for evaluating language development across the school years in Modern Israeli Hebrew, a language that shows marked distinctions in levels of usage between the everyday colloquial style and more formal types of expression due to its multilayered history (Shlesinger, 2000).[6] They showed that high register, semantically sophisticated, and often morphologically complex lexical items occurred with greater frequency in texts of older speaker/writers, especially in written expositions. Recently, Ravid and Vered (in press) showed in a structured elicitation task across adolescence that correct performance on neutral-register passive Hebrew verbs was significantly earlier than high-register passive verbs.

These findings have led us to adopt register as our independent measure as a proxy for frequency and familiarity assessment, hypothesizing that lower and neutral register adjectives should be more prevalent and thus learned earlier, while high register adjectives should be more typical of the lexicons of older, highly literate speaker/writers. We also assumed that we would find a relationship between the morphological structures of adjectives and their register ranking.

Method

The methodology of the current study consisted of two major steps: (i) compiling the adjective list, and (ii) eliciting the experts' register ranking judgements.

Step 1: Compiling the adjective list

All words marked as adjectives in Avneyon (2007) were identified, with a few newer adjectives (e.g., *viráli* 'viral') added from the media. Literary ad-hoc adjectives considered unfamiliar even to highly literate Hebrew users (e.g., *tafush* '(probably) stupid'), highly specific vocational loan adjectives (e.g., *osmofíli* 'osmophilic'), and highly ambiguous class items were excluded from the list. The master list of adjectives, our study corpus, contained 3,747 Hebrew adjectives. While we cannot claim that it contains all of the current adjectives in Hebrew, all items included in it are unambiguously Hebrew adjectives. The analyses below thus treat adjective lemma types (rather than wordform tokens occurring in discourse corpora), and for the current purposes they disregard inflectional adjective morphology, a topic that has gained its own set of developmental studies (Ravid & Schiff, 2012).

6. For example, 'tree' is denoted by both neutral register *ec,* deriving from Biblical Hebrew, and high-register *ilan*, from Mishnaic Hebrew.

Table 1. Morphological make-up of the adjective list

Category	Sub-category	Morphological structure	Example
Denominal		*i*-suffixed	*ma'asi* 'practical' [*ma'ase* 'action' + *i*]
Verbal patterns Present-tense *beynoni*	Resultative passives	*CaCuC*	*sagur* 'closed'
		meCuCaC	*mesudar* 'neat'
		muCCaC	*muklat* 'recorded'
	Non-passive participles	*CoCeC*	*bolet* 'prominent'
		CaCeC	*ayef* 'tired'
		niCCaC	*nifrad* 'separate'
		maCCiC	*mafxid* 'frightening'
		meCaCeC	*meza'azéa* 'shocking'
		mitCaCeC	*mictaber* 'compiled'
Nominal patterns		*CaCCan*	*aclan* 'lazy'
		CaCiC	*ta'im* 'tasty'
Non-productive	Handicapped	*CiCeC*	*iver* 'blind'
		CaCaC	*xaxam* 'wise'
	Incl. colors	*CaCoC*	*raxok* 'distant'
Diminutive	Reduplicative	*CCaCCaC*	*kxalxal* 'blueish'
Complex			*du-parcufi* 'two-faced'
Singleton			*meheyman* 'credible'
Miscellaneous slang			*ahbal* 'moron'

Morphological structures

Table 1 presents the morphological make-up of the adjectives in the master list, with examples. The lists consists of 19 different structures serving adjectives in nine major categories: (1) verb-based resultative passives; (2) verb-based non-passives; (3) nominal-based patterns; (4) nominal non-productive patterns; (5) denominal *-i* suffixed adjectives; (6) diminutive reduplicative adjectives; (7) complex and compound adjectives; (8) singletons; and (9) miscellaneous slang terms.

Step 2: Eliciting register judgments from language experts

Two pilot studies preceded the final elicitation of register judgment, designed to determine the expert judge population, the minimal number of expert judges necessary for inter-judge agreement, and questionnaire length. The pilots showed that native-speaking, well-educated Hebrew adults of different academic backgrounds could not reach adequate agreement on adjective register ranking. In contrast, native Hebrew speaking speech-language clinicians with linguistic and

psycholinguistic training were familiar with the notion of *language register,* and their responses showed much less variability, being more concentrated around a specific range of register ranks for most adjectives. The population selected as expert judges of language register rank was thus that of speech-language clinicians. The pilot studies also determined that the minimal number of judges per adjective was seven, and that 100 adjectives per questionnaire was an appropriate number for a single person to judge.

Participants and procedure

A set of 38 different questionnaires (37 containing 100 adjectives each, and one containing 50 adjectives) was compiled out of the randomized master list of 3,747 adjectives. They were disseminated electronically to 329 expert judges by email and via Facebook for volitional register judgment. Judges were instructed to intuitively classify the list of 100 adjectives into 5 register ranks, following examples from Rank 1 (basic register, e.g., *adom* 'red'), Rank 2 *barur* 'clear', Rank 3 *mevix* 'embarrassing', Rank 4 *meheyman* 'credible' and Rank 5 (very high register, *virtuózi* 'virtuose').

Data analysis

The 3,747 adjectives in 38 questionnaires were ranked by register rank on a scale of 1 to 5 by 329 language experts. Of these, 19 questionnaires were each ranked by seven expert judges, and 19 by more than seven experts, altogether yielding 32,329 rankings. Given the variance among experts' ranking of each adjective, among the rankings of each expert judge, and among all rankings, a cross-classification multi-level model was constructed to analyze the variance deriving from measurements in the different groups and levels (Level 1 – a single ranking of an adjective by an expert judge, Level 2 – expert judge, Level 2 – adjective). The mean of each adjective's ranking was calculated, as well as the mean for each expert judge's ranking. At Level 2, the mean distance of a single expert judge's ranking from his/her peers was calculated as a measure of expert judge agreement with peers. And at Level 2, the mean and variance values were calculated for each adjective across expert judges.

ICC analyses

Table 2 presents intra-class correlations (ICC), central tendency and dispersion values for experts and for adjective categories, so as to determine the multi-level structure of the rankings (over 0.05). The rightmost two columns present the two

Table 2. Intra-class correlations (ICC) for adjective categories, register ranking and variance

Morphological structure	Level 2 Adjectives	Level 2 Expert judges	# Adjectives in category	# Expert judges	# Rankings	Mean register ranking	Variance in ranking
1. *i*-suffixed	.49	.16	1,436	329	12,401	3.57	0.62
2. *CaCuC*	.54	.13	416	329	3,649	3.06	0.61
3. *meCuCaC*	.48	.16	603	329	5,176	3.09	0.57
4. *muCCaC*	.40	.23	178	321	1,582	3.02	0.61
5. *CoCeC*	.48	.17	94	298	805	2.65	0.62
6. *CaCeC*	.64	.14	56	243	446	2.48	0.60
7. *niCCaC*	.56	.09	68	278	572	3.17	0.51
8. *maCCiC*	.48	.18	93	304	786	2.70	0.59
9. *meCaCeC*	.45	.16	81	300	696	2.72	0.57
10. *mitCaCeC*	.32	.20	23	139	189	2.66	0.74
11. *CaCCan*	.41	.20	102	288	890	2.70	0.64
12. *CaCiC*	.56	.14	121	322	1,103	2.64	0.54
13. *CiCeC*	.64	.14	12	73	94	2.86	0.62
14. *CaCaC*	.69	.07	24	156	239	1.97	0.45
15. *CaCoC*	.68	.08	43	202	356	2.21	0.51
16. *CCaCCaC*	.46	.04	28	168	233	2.76	0.65
17. Complex	.30	.28	278	329	2,447	3.59	0.68
18. Singleton	.72	.08	55	253	491	2.41	0.50
19. Miscel.	.38	.33	36	219	325	2.36	1.05
Total	.42	.16	3,747	329	32,529	3.20	0.61

values central to this analysis for each of the structural categories: the mean register rank and the amount of variance among expert judges, so that the higher the variance, the less agreement among expert judges regarding register ranking. Following these analyses, 10 expert judges with consistently outlier rankings (two-sided confidence interval at a 95% level of confidence) were removed from the following analyses so as to achieve a classification of the adjectives into meaningful groups with low variance (= high agreement) among expert judges.

Latent Class Analyses

The second step of analysis was a Model Based Latent Class Analysis (LCA) that enables the identification of unobservable subgroups that are similar, based on observed characteristics – in the current case, mean register ranks and variance in ranking for each of the adjectives. The purpose of this analysis was to find the optimal number of register / variance groups in the data, unrelated to their morphological classification. The methodology developed to determine the optimal

Table 3. Comparing clustering in an ascending number of groups

	4 groups	5 groups	6 groups	7 groups	8 groups	9 groups	10 groups
AIC	12601	12342	12168	12070	11982	11899	11829
BIC	12682	12442	12286	12207	12138	12073	12022
Loglikelihood	–6287	–6155	–6065	–6013	–5966	–5921	–5883
Entropy	.73	.69	.73	.74	.75	.75	.78
# of adjectives							
1	1639	1779	1174	1046	1244	325	258
2	62	630	505	406	134	215	919
3	798	832	223	121	225	581	507
4	1248	452	1385	1246	340	924	256
5		54	404	214	396	389	215
6			56	683	31	805	687
7				31	320	31	341
8					1057	358	502
9						119	47
10							15
Vuong-lo-Mendell	< .001	.003	< .001	.03	< .001	.01	.23
Lo-Mendell-Rubin	< .001	.003	< .001	.03	< .001	.01	.24
Test for n-1 versus n clusters	< .001	< .001	< .001	< .001	< .001	< .001	< .001

Table 4. The mean probability of belonging to a group (in percentages) from 4 to 9 groups

	4 groups	5 groups	6 groups	7 groups	8 groups	9 groups
1	.87	.74	.80	.81	.79	.77
2	.94	.80	.86	.81	.83	.83
3	.82	.80	.84	.85	.88	.79
4	.83	.87	.77	.78	.73	.74
5		.95	.83	.81	.85	.80
6			.95	.81	.97	.77
7				.96	.71	.95
8					.82	.77
9						.85

number of groups was based on a multiple-test criterion for which the objective is to optimize classification goodness of fit subject to highest matching probability. This was done to ensure higher group-division of data having insignificant contribution to the goodness-of-fit (Lo-Mendell-Rubin test; Lo, Mendell, & Rubin, 2001), whereas items have a higher probability to belong in the assigned

group rather than in another (Nylund, Asparouchov, & Muthen, 2007). Table 3 shows that according to the goodness-of-fit criterion, the 10th group does not contribute additional explanation of the variance beyond nine groups (p = .24). This is also indicated by the AIC and BIC values (lower values indicate a better quality of the model) and the entropy value (the closer to 1, the better quality of group classification). Table 4 shows that the probability of an adjective belonging to a group is not lower than 80% up to 7 groups, decreasing beyond that. Taken together, a structure of seven latent groups was selected for our analysis.

Results

Quantitative analyses

Table 5 and Figure 1 show the grouping of all 3,747 adjectives on the master list into the seven groups emerging from the analysis described above. They depict the sizes of the adjective groups and their differing register ranks, ranging between a

Table 5. Number of adjectives, percentages out of the total number of adjectives (N = 3,747), mean register levels and variance in seven latent groups

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7
Mean register rank	2.49	4.51	3.18	3.71	1.44	3.12	3.08
Variance	0.49	0.22	1.58	0.52	0.21	1.03	2.46
N adjectives	1046	406	121	1246	214	683	31
Percentage	28%	11%	3%	33%	6%	18%	1%

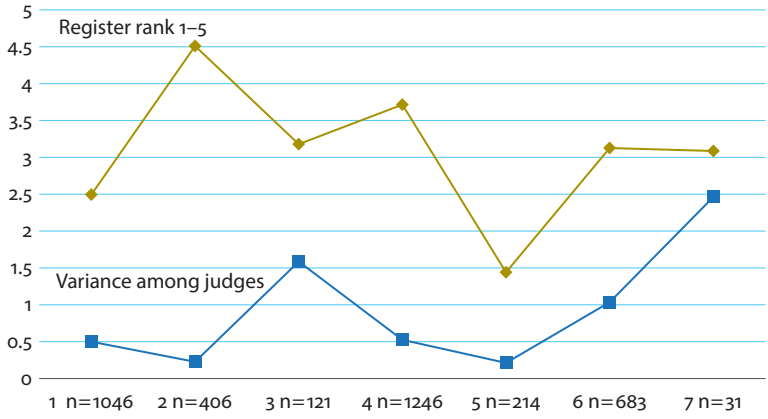

Figure 1. Seven latent group of register rank and expert judges' agreement

Table 6. The five adjective lexicons of Hebrew

Lexicon, percentage		Register	Examples
Core	6%	Basic	*xam* 'hot', *xadash* 'new', *rek* 'empty'
Basic	29%	Neutral	*ragzan* 'cranky', *zahir* 'careful', *ani* 'poor'
Mature	19%	Elevated	*megusham* 'awkward', *biológi* 'biological', *atum* 'opaque'
Educated	35%	High	*xamakmak* 'elusive', *dagul* 'eminent', *topográfi* 'topographic'
Super-literate	11%	Very high	*kamai* 'primeval', *menuce* 'feathered', *taktíli* 'tactile'

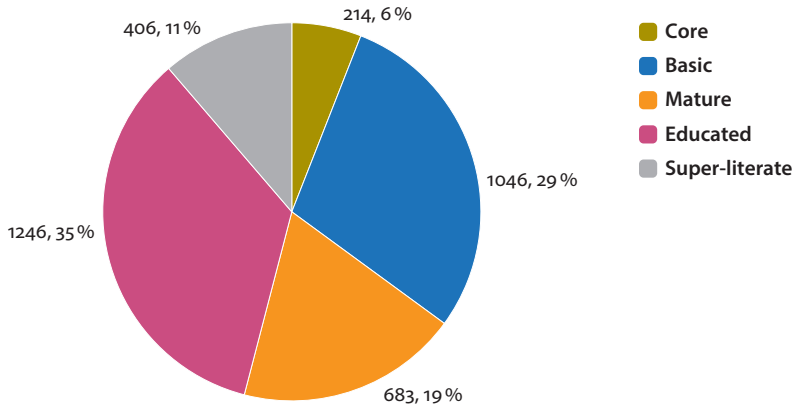

Figure 2. The five adjective lexicons of Hebrew, by linguistic register (N = 3,595)

low 1.44 to a high 4.51. Most important, they indicate that five of these groups – 1, 2, 4, 5, and 6 – had low expert judge variance values ranging between 0.21 to 1.03, where high variance across judges means lower agreement, that is, judges' evaluations is sparse around the mean value. These five groups showed middle to high agreement among expert judges. As adjectives come in groups, we decided to exclude small groups which also showed low agreement as expressed in the sparsity of the judges' evaluations

A final master list of 3,595 adjectives divided into five groups with high expert judge agreement and different register rankings constituted the focus of our further analyses. Given their register ranking and the lexical semantics of their adjective members, these five groups were taken to designate five developmentally consecutive "adjective lexicons". Table 6 and Figure 2 show the number and proportions of each lexicon out of the total of adjectives, with examples.

The five adjective lexicons: Semantic and morphological analyses

Below we present semantic and morphological analyses of the adjectives in each of the five lexicons. We based our semantic analyses on Blackwell's (2005) and Tribushinina et al.'s (2014) semantic coding of adjectives in CDS and CS corpora. However, as we show below, this classification mostly applied to the lower register lexicons, and did not cover the most prevalent abstract adjective classes. This is where morphological categorization made a contribution towards distinguishing the lexicons. Figures 3–7 depict the morphological makeup of the five adjective lexicons.

The Core Adjective Lexicon: Lowest register (M = 1.44)

This was the smallest lexicon, with 214 adjectives constituting 6% of the total of adjectives in the final master list, and the lowest mean register rank of 1.44. It contained what are universally regarded as the most basic adjectives in toddlers' and young children's perception of objects and people (Pitchford & Mullen, 2001), prevalent in their speech and in caregivers' child addressed speech (Blackwell, 2005; Tribushinina et al., 2014). Semantically, this list included everyday, mostly externally perceived, often concrete semantic categories of adjectives, each containing a small number of the most prominent adjectives in the category. Following Tribushinina et al.'s (2014) classification, these categories designated age (*xadash* 'new'), behavioral properties (*macxik* 'funny'), color terms (*adom* 'red'), conformity (*ragil* 'usual'), consistency (*rax* 'soft'), modals (*carix* 'necessary'), ordinal numbers (*rishon* 'first'), physical properties (*shaket* 'quiet'), physical states (*met* 'dead'), quantitative values (*rek* 'empty'), taste (*matok* 'sweet'), shape (*agol* 'round'), spatial values (*gavóha* 'tall'), and temperature (*xam* 'hot'). In addition, there were adjectives denoting basic internal states, sensations and emotions (*acuv* 'sad') and basic evaluative adjectives (*xamud* 'cute').

Morphologically, Figure 3 shows that the Core Adjective Lexicon contained the widest array of different structures. Most prominent (about ¼ of all adjectives in this lexicon) were the resultative *CaCuC* (*xamuc* 'bitter') and *meCuCaC* (*meluxlax* 'dirty), followed by about 10% *-i* suffixed adjectives, mostly simplex (*naki* 'clean'). Almost all non-passive *beynoni* participles were also represented in this lexicon, most prominently the *qal* allomorph *CaCeC* (*shamen* 'fat') and the *hif'il beynoni maCCiC* (*mafxid* 'scary'), in addition to *CaCoC* colors (*shaxor* 'black'), basic *CaCiC* adjectives (*ca'ir* 'young'), and less productive nominal patterns, e.g., *CaCaC* (*kacar* 'short'). While there were relatively numerous singletons (*kal* 'light'), this lexicon did not contain complex adjectives.

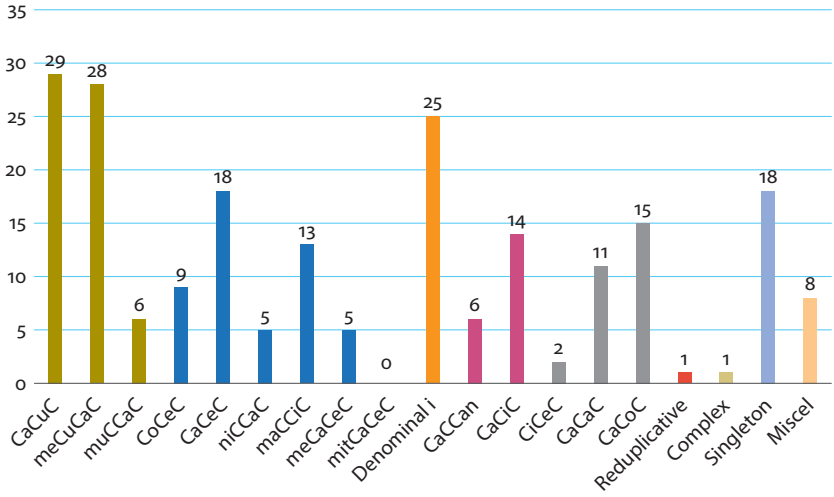

Figure 3. The Core Adjective Lexicon, by morphological category (N = 214)

The Basic Adjective Lexicon: Neutral register (M = 2.49)

This lexicon, scoring a mean of 2.49 out of 5 on register, was the second largest on the master list, containing 1046 (29%) of the adjectives. This was the adjective lexicon necessary to denote main and inherent properties of nouns, taken to represent the broad adjective lexicon shared by all Hebrew speakers with elementary education. There were far fewer concrete and externally perceived adjectives in this lexicon, relating to less prominent physical facets of objects (*zakuf* 'upright'), of the environment (*me'unan* 'cloudy', *mehavhev* 'blinking'), and of people (*mezukan* 'bearded'). Many of these external adjectives were lexically specific (*mezoham* 'filthy' rather than 'dirty', *zehe* 'identical' vs. *shone* 'different'). The majority of the Basic adjectives were abstract, expressing internal human attributes and states (*yediduti* 'friendly', *bayshan* 'shy'), advanced modals (*mesugal* 'able') and evaluations (*doxe* 'repulsive', *shaguy* 'mistaken', *yahir* 'arrogant'). Many adjectives expressed facets of objects and events that can only be captured by thought (*pit'omi* 'sudden', *meyaceg* 'representative', *mezuyaf* 'fake'), or require understanding of scientific principles (*nozli* 'liquid', *dalik* 'flammable', *sidrati* 'serial').

This lexical growth was accompanied by the concurrent changes in the morphological makeup of the Basic versus the Core Adjective lexicons (Figure 4). Whereas the Core Lexicon had 54% non-passive participles, nominal patterns and singletons, the Basic Lexicon had half that many: Non-passive *beynoni* participles (*ashem* 'guilty', *me'axzev* 'disappointing') took up about 16% of this lexicon, and nominal patterns denoting human qualities (*xashdan* 'suspicious', *zahir* 'careful') contributed another 10%, with about 1% singletons.

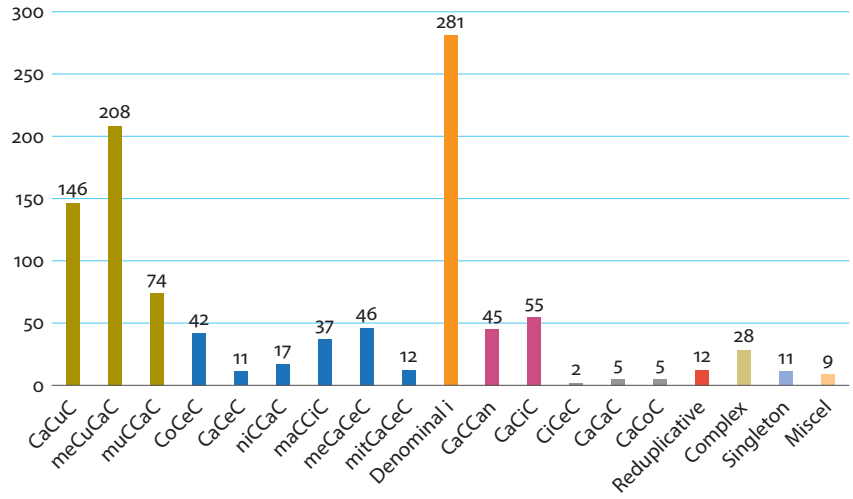

Figure 4. The Basic Adjective Lexicon, by morphological category (N = 1,046)

The Basic Lexicon was characterized by the two prominent morphological structures of Modern Hebrew. There was a big chunk (40%) of adjectives in verb-derived resultative passive *CaCuC, meCuCaC* and *MuCCaC*, expressing the *states* of objects, people, events and concepts (*katuv* 'written', *memushax* 'lasting', *mufla* 'magical'); and 25% denominal adjectives denoting the *attributes* of objects, people, events and concepts (*arci* 'earthy', *hitpatxuti* 'developmental', *mikco'i* 'vocational').

The Mature Adjective Lexicon: Elevated register (M = 3.12)

The Mature Adjective Lexicon, with a mean register rank of 3.12 out of 5, containing 683 (19%) adjectives, was a vocabulary detailed enough to express qualities and properties of humans, objects, places and concepts in the modern world, taken to serve literate Hebrew speakers with a highschool education.

Semantically, this lexicon had hardly any of the physical and external attributes found in the two lower register lexicons. It had a small number of adjectives that required familiarity with special properties of materials and objects, such as *mushxal* 'threaded' and *kalúa* 'woven'. Otherwise, this lexicon was all about abstract human states and attributes (*gluy lev* 'frank', *meyusar* 'tormented'), including qualities relevant to the adult world (*shtuf zima* 'salacious', *pasivi* 'passive'). Numerous adjectives testified to a systematic, school-based understanding of objects, systems, concepts and events (*xadish* 'modern', *ezraxi* 'civilian', *xofef* 'concurrent'). Again, this semantic shift was reflected in the morphological makeup of

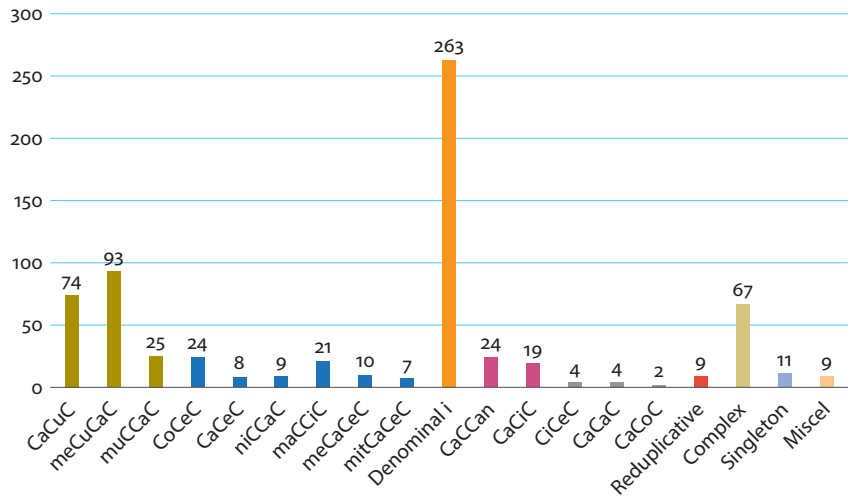

Figure 5. The Mature Adjective Lexicon, by morphological category (N = 683)

the Mature Lexicon (Figure 5). Close to 40% of the Mature adjectives were *i*-suffixed denominals (*xatuli* 'feline', *vatrani* 'lenient', *mishpati* 'legal') – among them, for the first time, foreign-based adjectives (*sensacióni* 'sensational', *globáli* 'global'); and another big chunk (28%) were passive resultatives. About 10% of this lexicon consisted of complex and compound adjectives (*du-mini 'bi-sexual', ktan koma* 'small (of) stature'). Under ¼ of the adjectives here belonged to non-passive participles and nominal patterns.

The Educated Adjective Lexicon: High register (M = 3.71)

This was the largest adjective lexicon, with a mean register rank of 3.71 out of 5, consisting of 1246 adjectives (35% of the final master list), a lexicon of lexically specific, morphologically complex adjectives serving educated Hebrew users. To gain command of the Educated Adjective Lexicon, one must have a rich nominal lexicon and well-developed derivational morphology abilities. Morphologically, its makeup resembled that of the Mature Lexicon, underscoring the composition of literate adjectives in Hebrew. The largest component was 42% denominal adjectives, many of them foreign-based (*vulkáni* 'volcanic', *ambiciózi* 'ambitious'), and native Hebrew (*trashi* 'lapidarian'), requiring familiarity with both the sciences and Classical Hebrew. A third of the Educated adjectives were passive resultatives (*mesuvag* 'classified', *gadúa* 'severed'), and about 11% were complex adjectives (*nesu panim* 'distinguished', *kcar ro'i* 'short sighted').

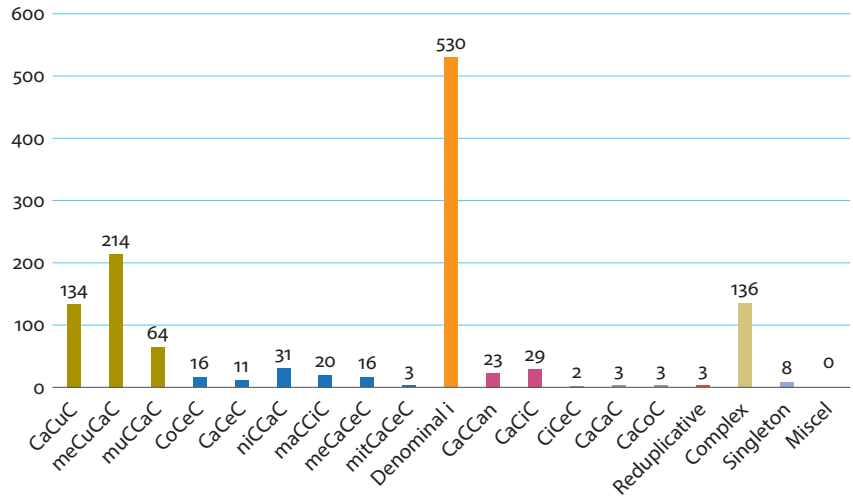

Figure 6. The Educated Adjective Lexicon, by morphological category (N = 1,246)

The Super-Literate Adjective Lexicon: Very high register (M = 4.51)

This second-smallest lexicon (406 adjectives, 11% of the total) with the highest mean register rank (4.51 out of 5) was taken to be the experts' adjective vocabulary. The Super-Literate Adjective Lexicon contained rare adjectives with high lexical specificity, serving highly educated Hebrew users with professional expertise. It required broad and deep conceptual knowledge, acquaintance with words

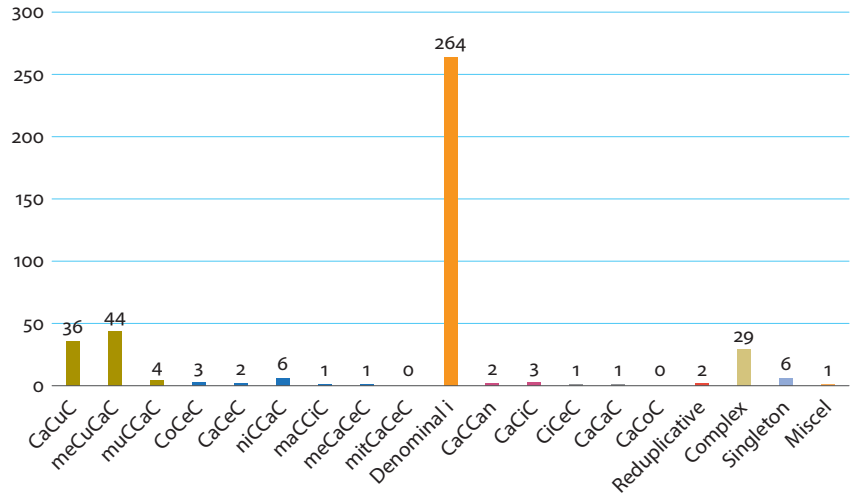

Figure 7. The Super-literate Adjective Lexicon, by morphological category (N = 406)

from Classical Hebrew, and terminology from foreign languages. Paradoxically, this Lexicon did not require much knowledge of morphology, as over 2/3 of it was composed of denominal adjectives, overwhelmingly foreign-derived (*kompulsívi* 'compulsive', *puritáni* 'puritan', and Hebrew *nici* 'hawkish'). About 20% were rare resultative passive adjectives such as *merubad* 'stratified' or *kahuy* 'torpid').

Discussion

The current study set out to propose subjective ranking by register as an independent measure of language knowledge across development. Carried out by language experts, with mostly high agreement among raters, a linguistic register scale from 1–5 by successfully distinguished five groups of over 3,500 Hebrew adjectives, which were respectively labeled the Core, Basic, Mature, Educated and Super-Literate Hebrew adjective lexicons.

The first major finding was to what extent language experts' judgments were in line with known attributes of language in acquisition. The judges were not asked about the estimated frequency of the adjectives on the list, nor about the assessed familiarity of Hebrew speakers of different age levels with them. Rather, the judges expressed their own perception about the contexts of usage of the ranked adjectives, implying their developmental characteristics (Ravid & Berman, 2009). Thus, most adjectives in the two lower register lexicons (*xazak* 'strong', *mo'il* 'helpful') were basic, inherent, and concrete, pertaining to a broad array of communicative contexts and intelocutors.[7] Increasing register rank in the next two lexicons indicated a growth in adjective abstractness and specificity (*karuy* 'mined', *xaluci* 'pioneering'), more restricted, literate, communicative settings, and knowledge of academic arenas, scientific topics, literature, history, and the arts. The lexicon with the highest register rank was mostly loan-based, requiring a high degree of linguistic and academic sophistication (*míti* 'mythical', *epidémi* 'epidemical'). Thus, while this study did not directly test the development of adjectives in Hebrew, it was developmentally-oriented, motivated and anchored in previous research, and yielded adjective lexicons that can be associated with the kind of Hebrew typically used at different age and cognitive levels. The measure used in achieving this classification fulfilled the requirements noted in the introduction by being independent of frequency and familiarity ratings.

7. A small-scale pilot at Tel Aviv University (Vichner, 2016) already indicates that the adjectives found in 30 storybooks targeted at children aged 1–8 years belong in the two lower register lexicons (Core and Basic).

Semantics

A second interesting finding related to the semantic changes characterizing the adjectives in the five consecutive lexicons, which revealed the need to go beyond early childhood in conducting AoA lexical studies. In the current study, the adjective categorization in Tribushinina et al. (2014) and Blackwell (2005) meaningfully classified the adjectives in the two lower register lexicons, but was not sufficient for capturing the qualitative distinctions among adjectives in the higher-rank lexicons. These advanced adjectives were either more lexically specific alternatives within the same categories (comparable, perhaps, to English *catlike* and *feline*), or else they fitted in the Tribushinina et al. category labeled "Other" (exemplified by *electric*). Thus, a more sophisticated classification, taking into account the cognitive, literacy and linguistic gains of adolescence is necessary to account for the changes in adjective semantics.

The lexical semantics of adjectives in the three higher register lexicons reflects a profound cognitive shift during the school years. The higher-register adjectives express scientific knowledge (*amórfi* 'amorphic'), sophisticated understanding of human relationships (*mefuyas* 'placated'), and the explosion of abstract nominals, encoding processes, activities, events, properties, states, concepts and ideas in literate expression, accompanied by adjectives denoting multiple facets of these abstract nominals (*dikduki* 'grammatical') (Nippold & Duthie, 2003; Ravid & Levie, 2010). These findings too underscore the need to further investigate language development beyond the early years (Berman, 2004; Nippold, 2016).

Morphology

A third, typologically important finding was the concurrent change in the distribution of Hebrew morphological categories side by side with the changes from general and concrete to specific and abstract semantics. The lower register lexicons, and in particular the Core lexicon, contained the fullest array of structural categories, including *beynoni* present participles in all *binyan* conjugations and all nominal patterns, including the least productive. The higher the register rank, the fewer the number of morphological categories in the lexicons – the two highest register lexicons contained mainly two, and then one structural device. We believe these differences reflect the interface of lexicon and morphology in the acquisition of the Hebrew adjective class, with the U-shape well known to researchers of language development. The small size of the Core lexicon and the small number of adjectives in each morphological category indicate that they are mostly learned as non-analyzed amalgams denoting primary and salient noun attributes in early

childhood. The proliferation of adjectives in the Basic lexicon (close to 1/3 of all adjectives on the master list) indicates the emergence of adjectival morphology in preschool and the early school years (Berman, 1985). In particular, this change points to the intensive learning of forms and meanings of resultative passive adjectives together with verb morphology (Berman, 1993, 1994; Ravid et al., 2016), and the inception of the class of denominal adjectives (Ravid & Nir, 2000). The decline in less productive adjectival forms and the increase in the highly productive resultatives and denominals in the Mature and Educated lexicons reflect the consolidation and automatization of Hebrew adjectival morphology as an accessible and flexible device during the school years (Ravid, 2004; Ravid & Zilberbuch, 2003). At the second end of the developmental U-shape, where mostly foreign-based denominal adjectives are learned as genre-related concepts and terms, morphology again plays a lesser role.

The analysis of the morphological makeup of the five adjective lexicons provides new information about the full developmental paths of adjective morphology. Modern Hebrew has come into its own in using two morphological devices for the expression of adjectival notions – *i*-suffixed denominals in the expression of nominal *attributes,* and verb-derived passive resultatives in the expression of nominal *states*.

The category of *i*-suffixed denominals presents clear linear development, increasing exponentially from 10% in the Core lexicon to two-thirds of the highest register lexicon. The findings of the current study thus support the centrality of denominal adjectives, the largest class in the three literate lexicons, as the main Hebrew mechanism conveying noun attributes, with specific syntactic affinity to abstract nominals (Ravid & Berman, 2010; Ravid & Levie, 2010; Ravid & Zilberbuch, 2003). The category undergoes its own evolution across the five lexicons from non-derived denominals in the Core lexicon (*acbani* 'nervous'), to Hebrew-based stems in the literate lexicons (*hatxalati* 'initial'), and finally to foreign stems in the Super Literate lexicon (*spiráli* 'spiral').

Resultative passive adjectives occupy at least 20% of all five lexicons, with a non-linear, protracted acquisition path. Previous studies pointed to the early childhood learning of resultative morphology (Berman, 1993, 1994), but the current study shows this category is in fact learned in two main waves. Resultative adjectives occupy ¼ of the core lexicon, increasing to 40% in the Basic lexicon as its largest category. This reflects the first wave of learning the present-tense forms of three different passive *binyan* conjugations to express the result of actions (e.g., *harus* 'destroyed' entailed by verb *haras* 'destroy', *mekupal* 'folded' entailed by *kipel* 'fold'). This proportion drops to 28% in the Mature lexicon, rising again to 33% in the Educated lexicon, reflecting a second wave of acquisition, with rarer and more passive-like resultatives coming in (*mushtal* 'implanted', *mekulas* 'eulogized')

(Ravid & Vered, in press). The smallest proportion of resultatives (20%) is found in the Super-literate lexicon. The internal evolution of this category in the two highest-register lexicons indicates the acquisition of flexible command of root and pattern morphology, with highest-register resultatives derived directly from nouns (and other bases) rather than verbs, e.g., *memutag* 'branded' from *mutag* 'brand', *menuce* 'feathered' from *noca* 'feather', *memusad* 'committed' from *mosad* 'institution'. Overall, the largest resultative class is *meCuCaC*, related to *pi'el,* the most productive *binyan* conjugation (Berman, 1993), followed by *CaCuC*, related to *qal,* the *binyan* with the largest number of verb types (Ravid et al., 2016), and finally *muCCaC*, the most passive-like resultative, related to *hif'il* (Ravid & Vered, in press).

Two other structures are of interest here. The first is the nominal pattern *CaCiC* (*yaxid* 'single') that actually hosts two different morphological entities and accordingly occupies different developmental time points. The first wave of *CaCiC* adjectives shares no specific adjectival meaning, unlike most productive noun and adjective patterns in Hebrew (Ravid, 1990). These mostly show up in the Core lexicon (*xarif* 'hot (taste)', *na'im* 'pleasant'). The second wave involves the morphologically complex potential adjectives (comparable to English *-able* adjectives), such as *amid* 'durable' or *pagía* 'vulnerable'. These occur with higher frequency in the Mature lexicon, and especially so in the morphologically complex Educated lexicon, associated with the high school and college years. This is the first empirical piece of evidence regarding the late development of the potential adjective category.

A second late occurring adjectival device labeled 'complex' adjectives, including adjective-headed compounds (*axuz dibuk* 'held (by the) devil = possessed') and prefixed adjectives (*rav-gili* 'multi-age'), showed up mostly in the Mature and Educated lexicons, providing support for previous findings (Ravid & Levie, 2010). Complex adjective numbers are probably greater than found here, as quantifier prefixes can create more than one complex adjective (e.g., *xad-sitri* 'one way', *du-sitri* 'two way'), and they also attach to free-standing adjectives (*du-cdadi* 'bi-directional'). This, however, would not change their high-register rating as part of the complex adjective lexicons of adolescent and adult Hebrew users.

Register as proxy for frequency

The subjective linguistic register rating of adjectives by language experts was successful as an independent proxy measure of lexical frequency, echoed in the concurrent evolution of adjective morphological devices. We now have in our possession five adjective lexicons arranged by developmental order, from which

the next stage of the project will draw the stimuli for AoA and clinical tasks. Although the efficacy of subjective register ranking was demonstrated here for Hebrew, a language with great historical depth and discrepancies between spoken and written styles (Ravid, 1995, 2012), we believe it can be adapted to other languages, subject to their typology and specific contexts of language variation. Three examples come to mind. For Arabic, a widely-spoken language with acute diglossia, written Standard Arabic and the spoken dialects would clearly frame the extreme poles of register, with criteria spanning the lexicon, morphology and syntax (Saiegh-Haddad & Spolsky, 2014). Although French diglossia is less well-recognized, spoken and written French morpho-syntax would be a good candidate for register distinctions (Gadet, 2009). And for English, the sweeping historical re-lexification responsible for its current double Latinate-Germanic lexicon has already been linked to register distinctions across development (Bar-Ilan & Berman, 2007).

References

Alderson, J. C. (2007). Judging the frequency of English words. *Applied Linguistics*, 28(3), 383–409. doi:10.1093/applin/amm024

Andersen, E.S. (1992). *Speaking with style: The sociolinguistic skills of children*. London: Routledge.

Auer, E. T., Bernstein, L. E., & Tucker, P. E. (2000). Is subjective word familiarity a meter of ambient language? A natural experiment on effects of perceptual experience. *Memory & Cognition*, 28, 789–797. doi:10.3758/BF03198414

Avneyon, E. (2007). *The Sapir Hebrew-Hebrew dictionary*. Or Yehuda: Hed Artzi. [in Hebrew]

Baayen, R. H., Feldman, L. B., & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language*, 55, 290–313. doi:10.1016/j.jml.2006.03.008

Balota, D. A., Pilotti, M., & Cortese, M. J. (2001). Subjective frequency estimates for 2,938 monosyllabic words. *Memory & Cognition*, 29, 639–647. doi:10.3758/BF03200465

Balota, D., Cortese, M., Sergent-Marshall, S., Spieler, D., & Yap, M. (2004). Visual word recognition for single-syllable words. *Journal of Experimental Psychology: General*, 133(2), 283–316. doi:10.1037/0096-3445.133.2.283

Bar-Ilan, L. & Berman, R. A. (2007). Developing register differentiation: The Latinate-Germanic divide in English. *Linguistics*, 45, 1–36. doi:10.1515/LING.2007.001

Bar-On, A., & Ravid, D. (2011). Morphological decoding in Hebrew pseudowords: A developmental study. *Applied Psycholinguistics*, 32, 553–581. doi:10.1017/S014271641100021X

Bar-On, A., Dattner, E., & Ravid, D. (Submitted). Resolving homography in reading Hebrew garden-path sentences: A developmental perspective.

Benelli, N., & Grinboim, Y. (2015). *Between familiarity and frequency: A psycholinguistic study of Hebrew adjectives*. Tel Aviv University ms. [in Hebrew]

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

Berman, R. A. (1985). *Acquisition of Hebrew*. Hillsdale, NJ: Lawrence Erlbaum.
Berman, R. A. (1993). Marking of verb transitivity by Hebrew-speaking children. *Journal of Child Language*, 20, 641–669. doi:10.1017/S0305000900008527
Berman, R.A. (1994). Formal, lexical, and semantic factors in acquisition of Hebrew resultative participles. *Berkeley Linguistic Society*, 20, 82–92. doi:10.3765/bls.v20i1.1452
Berman, R. A. (2004). Between emergence and mastery: The long developmental route of language acquisition. In R. A. Berman (Ed.), *Language development across childhood and adolescence: Psycholinguistic and crosslinguistic perspectives* (pp. 9–34). Amsterdam: John Benjamins. doi:10.1075/tilar.3.05ber
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. doi:10.1017/CBO9780511609664.007
Biber, D., & Conrad, S. (2009). *Register, genre, and style*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511814358
Blackwell, A. A. (2005). Acquiring the English adjective lexicon: Relationships with input properties and semantic typology. *Journal of Child Language*, 32, 535–562. doi:10.1017/S0305000905006938
Blodgett, E. G., & Cooper, E. B. (1988). Talking about it and doing it: Metalinguistic capacity and prosodic control in three to seven year olds. *Journal of Fluency Disorders*, 13, 283–290. doi:10.1016/0094-730X(88)90019-8
Brysbaert, M., & Cortese, M. (2011). Do the effects of subjective frequency and age of acquisition survive better word frequency norms? *The Quarterly Journal of Experimental Psychology*, 64(3), 545–559. doi:10.1080/17470218.2010.503374
Brysbaert, M., & Ghyselinck, M. (2006). The effect of age of acquisition: Partly frequency related, partly frequency independent. *Visual Cognition*, 13, 992–1011. doi:10.1080/13506280544000165
Carroll, J. B. (1970). An alternative to Juilland's usage coefficient for lexical frequencies. *ETS Research Bulletin Series*, 1970(2), i–15.
Caselli, M. C., Bates E., Casadio, P., Fenson, J., Fenson, L., Sanderl, L., & Weir, J. (1995). A cross-linguistic study of early lexical development. *Cognitive Development*, 10, 159–199. doi:10.1016/0885-2014(95)90008-X
Clark, E.V. (2003). *First language acquisition*. Cambridge: Cambridge University Press.
Diessel, H. (2007). Frequency effects in language acquisition, language use, and diachronic change. *New Ideas in Psychology*, 25, 108–127. doi:10.1016/j.newideapsych.2007.02.002
Dixon, R. M. W., & Aikhenvald, A. Y. (2014). *Adjective classes: A cross-linguistic typology*. Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780198701316.001.0001
Dockrell, J. E., & Messer, D. (2004). Later vocabulary acquisition. In R. A. Berman (Ed.), *Language development across childhood and adolescence: Psycholinguistic and crosslinguistic perspectives* (pp. 35–52). Amsterdam: John Benjamins. doi:10.1075/tilar.3.06doc
Douglas, F. M. (2003). The Scottish corpus of texts and speech: Problems of corpus design. *Literary and linguistic computing*, 18(1), 23–37. doi:10.1093/llc/18.1.23
Ferguson, C. A. (1994). Dialect, register, and genre: Working assumptions about conventionalization. In D. Biber & E. Finegan (Eds.), *Sociolinguistic perspectives on register* (pp. 15–30). New York: Oxford University Press.
Gai, A. (1995). The category 'adjective' in Semitic languages. *Journal of Semitic Studies*, 1, 1–9. doi:10.1093/jss/XL.1.1

Gadet, F. (2009). Stylistic and syntactic variation: Introduction. In K. Beeching, N. Armstrong, & F. Gadet (Eds.) *Sociolinguistic variation in contemporary French* (pp. 115–120). Amsterdam: John Benjamins. doi:10.1075/impact.26.09gad

Gries, S. T. & Divjak, D. (Eds.) (2012). *Frequency effects in language learning and processing*. Berlin: De Gruyter Mouton. doi:10.1515/9783110274059

Hora, A., Avivi-Ben Zvi, G., Levie, R., & Ravid, D. (2006). Acquiring diminutive structures and meanings in Hebrew: an experimental study. In I. Savickiene & W. U. Dressler (Eds.), *The acquisition of diminutives* (pp. 295–317). Amsterdam: John Benjamins. doi:10.1075/lald.43.13hor

Itai, A., & Wintner, S. (2008). Language resources for Hebrew. *Language Resources and Evaluation*, 42, 75–98. doi:10.1007/s10579-007-9050-8

Juhasz, B. J. (2005). Age-of-acquisition effects in word and picture identification. *Psychological Bulletin*, 131, 684–712. doi:10.1037/0033-2909.131.5.684

Jurafsky, D. (2003). Probabilistic modeling in psycholinguistics: Linguistic comprehension and production. In R. Bod, J. Hay, & S. Jannedy (Eds.), *Probability theory in linguistics* (pp. 39–96). Cambridge, MA: The MIT Press.

Kennedy, G. (2014). *An introduction to corpus linguistics*. London & New York: Routledge.

Kuperman, V., & Van Dyke, J. A. (2013). Reassessing word frequency as a determinant of word recognition for skilled and unskilled readers. *Journal of Experimental Psychology: Human Performance and Perception*, 39, 802–823. doi:10.1037/a0030859

Levin, I., Ravid, D., & Rappaport, S. (2001). Morphology and spelling among Hebrew-speaking children: From kindergarten to first grade. *Journal of Child Language*, 28, 741–769. doi:10.1017/S0305000901004834

Lo, Y., Mendell, N., & Rubin, D. (2001). Testing the number of components in a normal mixture. *Biometrika*, 88, 767–778. doi:10.1093/biomet/88.3.767

MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk. Vol. 2: The database* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.

Nippold, M. A. (2016). *Later language development: School-age children, adolescents, and young adults* (4th ed.). Austin, TX: Pro-Ed.

Nippold, M. A., & Duthie, J. K. (2003). Mental imagery and idiom comprehension: A comparison of school-age children and adults. *Journal of Speech, Language, and Hearing Research*, 46, 788–799. doi:10.1044/1092-4388(2003/062)

Nylund, K. L., Asparouhov, T. & Muthen, B. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation. *Structural Equation Modeling*, 14, 535–569. doi:10.1080/10705510701575396

Pitchford, N. J. & Mullen, K. T. (2001). Conceptualization of perceptual attributes: A special case for color? *Journal of Experimental Child Psychology*, 80, 289–314. doi:10.1006/jecp.2001.2634

Ravid, D. (1990). Internal structure constraints on new-word formation devices in Modern Hebrew. *Folia Linguistica*, 24, 289–346. doi:10.1515/flin.1990.24.3-4.289

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

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: John Benjamins. doi:10.1075/lald.28.14rav

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: John Benjamins. doi:10.1075/tilar.3.07rav

Ravid, D. (2006). Word-level morphology: A psycholinguistic perspective on linear formation in Hebrew nominals. *Morphology*, 16, 127–148. doi:10.1007/s11525-006-0006-2

Ravid, D. (2012). *Spelling morphology: The psycholinguistics of Hebrew spelling*. New York: Springer. doi:10.1007/978-1-4419-0588-8

Ravid, D., & Berman, R. A. (2009). Developing linguistic register across text types: The case of Modern Hebrew. *Pragmatics & Cognition*, 17, 108–145. doi:10.1075/pc.17.1.04rav

Ravid, D. & R. Berman. (2010). Developing noun phrase complexity at school-age: A text-embedded cross-linguistic analysis. *First Language*, 30, 3–26. doi:10.1177/0142723709350531

Ravid, D., & Levie, R. (2010). Hebrew adjectives in later language text production. *First Language*, 30, 27–55. doi:10.1177/0142723709350529

Ravid, D., & Nir, M. (2000). On the development of the category of adjective in Hebrew. In M. Beers, B. van den Bogaerde, G. Bol, J. de Jong, & C. Rooijmans (Eds.), *From sound to sentence: Studies on first language acquisition* (pp. 113–124). Groningen: Center for Language and Cognition.

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

Ravid, D. & R. Schiff. (2012). From dichotomy to divergence: Number/gender marking on Hebrew nouns and adjectives across schoolage. *Language Learning*, 62, 133–169. doi:10.1111/j.1467-9922.2011.00675.x

Ravid, D., & Shlesinger, Y. (1987). On the classification and structure of *-i* suffixed adjectives. *Hebrew Linguistics*, 25, 59–70. [in Hebrew]

Ravid, D., & Vered, L. (In press). Learning passive forms as lexical items: Later language development in Hebrew. *Journal of Child Language*.

Ravid, D., & Zilberbuch, S. (2003). The development of complex nominals in expert and non-expert writing: A comparative study. *Pragmatics and Cognition*, 11, 267–297. doi:10.1075/pc.11.2.05rav

Ravid, D., Levie, R., & Avivi-Ben Zvi, G. (2003). Morphological disorders. In L. Verhoeven & H. van Balkom (Eds.), *Classification of developmental language disorders: Theoretical issues and clinical implications* (pp. 235–260). Mahwah, NJ: Erlbaum.

Ravid, D., Ashkenazi, O., Levie, R., Ben Zadok, G., Grunwald, T., Bratslavsky, R., & Gillis, S. (2016). 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* (pp. 95–134). Amsterdam: John Benjamins. doi:10.1075/tilar.19.04rav

Saiegh-Haddad, E. & Spolsky, B. (2014). Acquiring literacy in a diglossic context: Problems and prospects. In Saiegh-Haddad, E. & Joshi, M. (Eds.) *Handbook of Arabic literacy: Insights and perspectives* (pp. 225–240). Dordrecht: Springer. doi:10.1007/978-94-017-8545-7_10

Salerni, N., Assanelli, A., D'Odorico, L., & Rossi, G. (2007). Qualitative aspects of productive vocabulary at the 200- and 500-word stages: A comparison between spontaneous speech and parental report data. *First Language*, 27, 75–87. doi:10.1177/0142723707067545

Sandhofer, C., Smith, L. B., & Luo, J. (2000). Counting nouns and verbs in the input: Differential frequencies, different kinds of learning? *Journal of Child Language*, 27, 561–585.

Schachter, P., & Shopen, T. (2007). Parts-of-speech systems. In T. Shopen (Ed.), *Language typology and syntactic description: Clause structure* (2nd ed., vol. 1, pp. 1–60). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511619427.001

Shlesinger, Y. (2000). *Journalistic Hebrew: Stylistic aspects of Israeli newspaper sections*. Beer-Sheva: Ben-Gurion University of the Negev Press. [in Hebrew]

Spencer, A. (1991). *Morphological theory*. Oxford, UK: Blackwell.

Thompson, G. L., & Desrochers, A. (2009). Corroborating biased indicators: Global and local agreement among objective and subjective estimates of printed word frequency. *Behavior Research Methods*, 41, 452–471. doi:10.3758/BRM.41.2.452

Tribushinina, E., Van den Bergh, H., Ravid, D., Aksu-Koç, A., Kilani-Schoch, M., Korecky-Kröll, K., Leibovitch-Cohen, I., Laaha, S., Nir, B., Dressler W. U., & Gillis, S. (2014). The first year of adjectives: A growth curve analysis of child speech and parental input. *Language, Interaction and Acquisition*, 5, 185–226. doi:10.1075/lia.5.2.02tri

Vichner, A. (2016). *Adjectives in young children's storybooks*. Tel Aviv University ms. [in Hebrew]

Corresponding address

Dorit Ravid
School of Education
Tel Aviv University
Tel Aviv 69978
Israel

doritr@post.tau.ac.il