

Nominals in expert and non-expert writing

Revised version June 7, 2003

Dorit Ravid and Shoshana Zilberbuch
Tel Aviv University

**The development of complex nominals
in expert and non-expert writing:
A comparative study**

Running title: Nominals in expert and non-expert writing

Address for Correspondence

Dr. Dorit Ravid

School of Education and the Dept. of Communications Disorders

Tel Aviv University

Tel Aviv 69978

ISRAEL

Fax: 972 3 5360394

Email: doritr@ccsg.tau.ac.il

Ravid, D. & S. Zilberbuch. The development of complex nominals in expert and non-expert writing: A comparative study. *Pragmatics and Cognition* 11.2, 267-297, 2003.

Abstract

This study examines the distribution of complex nominal constructions in Hebrew texts produced by non-expert schoolage and adult writers, compared with their distribution in expert-written encyclopedic texts. One aim of the paper was to determine young writers' ability to distinguish text types through their usage of genre-appropriate morpho-syntactic forms. Another aim was to investigate the distribution of these constructions in expert school-related texts so as to confirm or refute the hypothesis of "resonance" between input and output texts. The study population consisted of native Hebrew-speaking gradeschoolers, adolescents and adults who each produced 2 written texts, one biographical, one expository. The same text types were analyzed in three encyclopedias, targeted for use in each of the age groups under discussion. Our analysis indicates that non-expert writers' textual output undergoes marked developmental changes during later childhood and adolescence, but no developmental resonance was found in expert texts, which constitute part of the input that novice and non-expert writers are exposed to. Implications for the nature of the mental lexicon and its relationship with grammar are discussed.

1.0 Introduction: Extending the framework

The study of language development has been extended in the last decades beyond its previous scope of investigation along four dimensions: age bracket, linguistic domains, modality, and genre. From *a developmental perspective*, it is clear by now that natural language acquisition is a protracted process that is not complete even by adolescence (Anglin 1993, Berman 1997, Nippold 1998). Most researchers would agree that children growing up in a monolingual environment have access to the vast majority of morphological and syntactic structures of their language before they reach school age; nonetheless, a five-year-old hardly matches an adult or even a twelve-year-old in linguistic proficiency. Besides increasing vocabulary, one significant aspect of later language development is the ability to recruit different morpho-syntactic resources and to use them flexibly for diverse communicative purposes (Berman in press a, Ravid & Tolchinsky 2002). This paper is concerned with two later-developing structures in the language of Hebrew-speaking gradeschoolers, adolescents and adults.

The *linguistic scope* of child language research has also been expanded considerably beyond the domains of lexicon and morpho-syntax to include the study of text production (Berman & Slobin 1994, Hickmann 1995). Two motivations underlie this shift. One is the fact that most marked, literate lexical items and morpho-syntactic structures do not occur in interactive conversation, the most “natural” form of language (Chafe 1994); however, they do occur in extended discourse and serve textual functions in specific text types encountered in the course of formal education (Berman in press b, Ravid in press, Ravid & Shlesinger 1995). For example, in analyzing the discourse role of nominalizations in English scientific texts, Halliday (1988) claims that “the device of nominalizing... is an essential

Nominals in expert and non-expert writing⁴

resource for constructing scientific discourse” (p.169) and has “two discursual effects: (...) packaging a complex phenomenon into a single semiotic entity, by making it one element of clause structure, so that (...) its rhetorical function - its place in the unfolding argument - is rendered fully explicit” (p. 168). In order to test children’s knowledge of morpho-syntactic constructions fulfilling textual functions it is necessary to look at their patterns of occurrence and distribution in authentic texts produced by language users of various ages.

Another motivation for looking at text production in a developmental perspective relates to the nature of acquisition. What does it mean for a structure to be acquired? Apart from the well-known differences between comprehension and production, it is not at all clear what is meant by the fact that a child performs successfully on an experimental test in which a construction is presented stripped of its context. Good performance may indicate that the child is familiar with said construction and has relevant syntactic and morpho-phonological *knowledge* of it, but does it mean that the construction has been acquired? Not until such constructions occur within the appropriate textual context can this question be answered positively. To take a specific example, Hebrew-speaking first graders perform fairly well on inflecting nouns with optional genitive suffixes (Levin, Ravid & Rappaport 2001), but analyses of oral and written texts produced by gradeschoolers, highschoolers and adults show that it is only the latter who are able to use optional bound morphology appropriately in discourse (Berman & Ravid 2003, Cahana-Amitay & Ravid 2000).

Modality – the physical mode of language production – is a third dimension in the newly developed framework for studying later language development. While the primary or unmarked form of language production is oral

interaction, there are other systems in which human language is encoded, such as written language. Language production is shaped in each modality under constraints and principles of human information processing such as speed, clarity, economy, and expressiveness (Clark 1996, Slobin 1977, Strömqvist & Wengelin 1999).

Written language is more planned, organized and coherent than spoken language, and the circumstances of its production encourage revision, review and rewriting, allowing the retrieval of higher-register, literate lexical items and marked morpho-syntactic structures during writing. The stability of the written text permits the production of longer and denser information packages in hierarchically complex syntactic constructions, which can be re-read and re-analyzed by readers without the pressures of on-line processing (Chafe 1994, Chafe & Danielewicz 1987). Learning to write in a sense means learning to deal with a new syntax, with producing and organizing larger linguistic units with the view of the text as a whole. Writing thus involves learning how to control and shape the flow of information in texts through linguistic means (Olson 1994, Ong 1992). Any investigation of later language acquisition in children should take the factor of modality into account, since it cannot be claimed that children are or are not familiar with a given construction without having given them the opportunity to use it in their writing (Pontecorvo 1997, Schneuwly 1997).

Genre. Texts are not uniform, and they vary across the dimension of *genre*, which is essentially textual in nature. Genres range across text types broadly defined by function, social-cultural practices, and communicative purpose - conversation, narration and exposition, information and poetry. A changing, open-ended complex of historical, social, cultural and communicative factors provide the frame for an array of specific text types or subgenres from commercials and contracts to

scientific writing and textbooks (Paltridge 1997). Mature, literate writers contextualize their text production in genre-appropriate forms, and genre-specific texts differ in thematic content, global structure, as well as in privileged rhetorical and grammatical constructions (Berman in press a,b). Different genres make their own demands regarding the selection and organization of thematic content and of privileged morpho-syntactic structures expressing this content. A major genre distinction is found between narratives, which focus on people, their actions and motivations, and express the unfolding of events in a temporal framework; and expository texts, which focus on ideas and concepts, and express the unfolding of claims and argumentation in a causal context (Berman & Katzenberger in press).

There is some evidence that gradeschool children are sensitive to the different textual character of narratives and expository texts. As early as in 4th grade, children make initial attempts at selecting phrasal, clausal and inter-clausal patterns that would be genre-appropriate (Berman & Ravid 1999; Ravid in press). Later-acquired morpho-syntactic devices such as *passé simple* in French narratives or passive constructions in Hebrew expository texts have been shown to serve genre-specific textual functions and to be diagnostic of age and modality (Berman & Ravid 1999; Gayraud, 2000). In this paper we proceed in the same direction, analyzing children's writing with the view to determine how they use grammatical constructions in their textual contexts.

1.1 Morpho-syntactic devices in later language development

Analysis of morpho-syntactic constructions within their textual context may provide information about later linguistic development at two levels. Firstly, studying the usage of target constructions in authentic texts written by children and adolescents will demonstrate to what extent they are able to go beyond the merely "correct" yet

decontextualized morpho-phonological production of these constructions to their appropriate usage in context. Secondly, this is a way to determine young writers' ability to distinguish text types and modalities through their usage of genre-appropriate morpho-syntactic forms. This is one goal of the current paper.

One domain which seems promising for the interface of morpho-syntax and text in development is the syntactic phrase. Clauses consist of syntactic phrases, specifically the predicate VP and satellite NPs. These are *bottom-up categories* whose internal structure, on the one hand, and the patterning of their combination with other phrases, on the other, produce differential clause-internal configurations. Phrases develop in length, complexity, and internal architecture with age and schooling (Scott 1988). Studies which trace the development of VPs, APs and NPs indicate that increasing complexity is expressed in choice of head type, in longer and less linear structures, in various types of conjoining and subordination within the phrase, and in a diversification of phrase structures (Ravid in press, Ravid & Avidor 1998, Ravid & Nir 2000). In this paper we focus on the internal structure of NPs, which permit an array of internal structures varying in complexity and configuration types. Initial studies have shown that written language, and especially expository texts, are characterized by complex lexical NPs which are crucial for the expression of hierarchical information structures participating in rhetorical text organization (Ravid, van Hell, Rosado & Zamora 2002).

1.2 Input and output in text development

A related issue to the question of children's production (or output) of later-acquired constructions is the problem of input. Major changes occur in children's lexical and syntactic knowledge during gradeschool (Karmiloff-Smith 1986, Levin et al. 2001, Scott 1988) as well as in their ability to perceive and recruit morpho-syntactic and

rhetorical structures in texts (Berman in press a, Berman & Katzenberger in press). Since the treatment of genre and modality in school teaching is highly implicit (Kress 1994), this development is most probably not the result of direct instruction. The major source of marked, literate lexical items and morpho-syntactic devices is exposure to written school-related language produced by expert writers (Anglin 1993, Nippold 1998, Smith 1998). Using school textbooks, encyclopedias, reference materials, dictionaries and Internet resources involves interpreting and internalizing genre-specific constructions in their authentic contexts. It can be assumed that school-related texts produced by expert writers are more fluently written and at a higher register than non-expert texts, including appropriate usage of complex syntactic constructions. It can be further assumed that expert writers should be able and willing to adjust their style and language to the level of their readers, that is, to use less dense complex constructions in texts for younger age groups, and to gradually raise linguistic sophistication in texts targeting older age groups. In other words, we can hypothesize a certain resonance between input and output texts. An investigation of the distribution of the same late-occurring constructions in expert school-related texts as in texts produced by children and adolescents might confirm or refute this hypothesis and yield further insights on later language acquisition. This is another goal of the current paper.

In this paper we examine the distribution of two types of complex nominal constructions in texts produced by non-expert schoolage and adult writers, compared with their distribution in expert-written, school-related texts targeting the age and schooling brackets that we are interested in. The next section presents the constructions under examination.

1.3 Complex nominals and their acquisition in Hebrew

Complex nominal constructions in Modern Hebrew constitute a wide range of compounding devices of various types, N-A and N-PP constructions, and nouns with relative clause and other sentential complementation (Berman 1988, Ravid & Shlesinger 1995). This paper examines two nominal constructions: N-N compounds and constructions with *i*-suffixed denominal adjectives (henceforth, A_{denS}).

1.3.1 *N-N compounds*. Compounds are critical to the syntax / lexicon interface since they allow speakers to combine words that express complex meanings transparently and productively, yet also have derivational characteristics that allow them to lexicalize easily as root compounds *deadline*, *diehard* and verbal compounds *rainmaker* and *graduation exercises* (Anderson 1985, Borer 1988, Lieber 1983). Hebrew compounds have traditionally been restricted to bound N-N head + adjunct structures, termed *smixut* 'adjacency', e.g., *sixat[^]télefon* 'conversation[^]telephone = telephone conversation'. Since the *smixut* head is bound to its modifier (signified by the caret sign), it may undergo morpho-phonological modifications (e.g. free stem *sixa* → bound stem *sixat*). As in other languages, compounds provide Hebrew speakers with a means of complex sub-categorization to a specific head (Berman 1988, Borer 1988), specifying a member of a head category as in English *telephone operator*.

N-N compounds occur at all levels of spoken and Hebrew, and constitute approximately one third of the nominal inventory of Modern Hebrew (Choueka 1996). Their proportion is a steady 20% out of all NPs across texts in various genres and modalities (Berman 1988, Ravid & Shlesinger 1995). However, the compounding device is not a preferred option among young Hebrew-speaking children. While young English speakers aged 3-4 would first use compounding to make up innovative labels such as *wagon man*, Hebrew speakers of the same age prefer word-internal morphology for lexical innovation, e.g., *agalan* 'wagoneer' (Clark & Berman 1984). By the time

Hebrew-speaking children have effectively mastered clause-internal syntax and inflectional morphology as well as a considerable amount of derivational morphology, compounding is still not a real option for them.

Compounding emerges as a productive lexical device in child Hebrew from age four, and from the beginning it is used exclusively for sub-categorization to designate contrasting sub-categories such as *tea pot* versus *coffee pot* (Berman 1987, Clark & Berman 1984). Sub-categorizing compounds also constitute the majority of N-N compounds across most text types in adult Hebrew (Ravid & Shlesinger 1995). No study has so far examined the distribution of N-N compounds in texts produced by older children and adolescents.

1.3.2 *Denominal adjective constructions*. Denominal *i*-suffixed adjectives (A_{denS}) are the most productive adjectival class in Modern Hebrew, e.g., *xašmal-i* ‘electr-ic’, *cibur-i* ‘publ-ic’. Structurally, A_{denS} involve the linear formation of a nominal stem (e.g., *xašmal* ‘electricity’) and an adjectival suffix *-i*. In many cases, the nominal stem undergoes morpho-phonological changes as a result of suffixation, e.g., *báyit* ‘house’ → *beyt-i* ‘domestic’. The semantics of A_{denS} is quite complex: In order to form a denominal adjective, the base noun has to be “dissolved” into its component properties to select the specific property that will be carried over to the derived adjective. The appropriate nominal property is not always predictable, as in other cases of denominal derivation (Aronoff 1980, Clark & Clark 1979).

Like all adjectives, A_{denS} have two syntactic functions: Predicative adjectives serving as predicate heads, and attributive adjectives with the function of NP modification. For example, the A_{den} *iši* ‘personal’ from *iš* ‘man, person’ may function as the predicate in *ha-sipur hu iši* ‘the letter is personal’ or as an NP modifier in *sipur iši* ‘story personal = personal letter’. In both cases, the adjective denotes a property

attributed to a noun - either the subject of the sentence or the NP head (Ferris 1993). Of the two syntactic positions, that of NP modification in an N-A_{den} construction is optional, carries extra information and involves more complex syntax than the obligatory position of an adjective as a predicate head, and accordingly it is also acquired later on (Ravid & Nir 2000, Shlesinger & Ravid 2003). The distribution of denominal adjectives in these two syntactic functions is thus expected to be diagnostic of age, modality, and genre, as outlined below.

N-A_{den} constructions, and especially those in the NP-modifier function, are typical of higher-register, written Hebrew, such as literary prose, journalistic and expository texts (Ravid & Shlesinger 1987). Apart from lexicalized forms such as *xagigi* 'festive', they are completely absent from child-directed speech. Accordingly, denominal *i*-suffixed adjectives are the last type of adjectives to emerge in Hebrew child language around age 6, often ill-formed, e.g., *xatani* 'groom-like' (Ravid & Nir 2000). Thus, *i*-suffixed denominal adjectives are a later acquisition than N-N compounds, and especially when occurring in a NP expressing complex sub-categorization. So far, no study has looked at the distribution of denominal adjectives in texts produced by children and adolescents.

1.3.3 *Compounds and N-A_{den}s as competing constructions.* N-A_{den}s constitute an alternative subcategorizing device to N-N compounds in Hebrew. The general semantic profile of N-A_{den}s is usually 'N₁ with the property of N₂', where the second N serves as the basis for denominal adjective formation, e.g., *sipur iši* 'story personal = personal story' from *iš* 'man, person'. Like N-N compounds, N-A_{den}s form complex nominals productively, yet lexicalize easily, their components blending to signify a semantically complex concept characterized by the idiosyncrasy and unpredictability of derivation (Aronoff 1976). N-A_{den} constructions may lexicalize to the point they

Nominals in expert and non-expert writing²

resemble exocentric compounds which refer to an external head, e.g., *roš turki* ‘Turkish head = a type of knot’.

N-N compounds and N-A_{den}s compete in the expression of head-modifier relations. They often interact in interesting, unpredictable and idiosyncratic ways (Berman 1988). There are cases where only one construction is possible (e.g., N-A_{den} *manhig ruxani* ‘spiritual leader’ but not compound *manhig rúax*; or compound *maxalat rúax* ‘mental disease’ but not N-A_{den} *maxala ruxanit*). In other cases the two constructions alternate with hardly any semantic distinction, e.g., compound *sxiyat láyla* ‘night swimming’ / N-A_{den} *sxiya leylit* ‘nocturnal swimming’.

Two studies have looked at the acquisition of these two constructions in Hebrew speakers. Ravid (1997) compared young children’s ability to produce N-N compounds and N-A_{den}s from kindergarten to first grade. This analysis showed that morpho-phonological knowledge of both types of constructions increased to the same extent during this period, however both age groups showed a better semantic grasp of compounding than of N-A_{den} formation. In a second study (Ravid & Zilberbuch 2003) of oral and written text production in Israeli gradeschoolers and highschoolers, N-N compounds were found to be prevalent in all text types and in both modalities: 90% of all complex sub-categorizing NP constructions in the gradeschoolers’ texts were N-N compounds, while in highschool this proportion declines to less than 75%, similar to expert texts. In contrast, A_{den}s appeared in all four text types only in highschoolers and above: Most 6th graders failed to produce any A_{den}s in some or most of their texts, and it was only in highschool that most texts contained at least one A_{den}. The number of both N-N compounds and A_{den}s increased with age, especially in the written modality.

1.3.4 Words and phrases in the mental lexicon.

Studying N-N compounds and N-A_{den}s provides an opportunity to examine the changing structure of the mental lexicon across development. Chomskyan linguistics makes a sharp distinction between generative syntax, supposedly the primary language component in human cognition, and a secondary rote-learned component, the lexicon (Chomsky 1970). But the study of real-world discourse in a developmental perspective indicates that this distinction is often blurred at a number of levels. First, the lexicon is an absolute necessity for the development of syntax: All syntactic phrases are crucially dependent on the occurrence of content words (Cahana-Amitay & Ravid 2000). Secondly, the higher the occurrence of content words, especially nouns and adjectives, the higher the complexity of syntactic architecture which frames them (Ravid in press). Thirdly, the very distinction between words and syntactic units falls apart under close examination, which reveals that idiomatic constructions held in memory range from morphemes and traditional ‘words’ to syntactic phrases and whole clauses (Jackendoff 2002). N-N and N-A_{den}s are not words and lend themselves to syntactic expansion, but note, for example, that an N-A_{den} expression such as *eš yedidutit* ‘friendly fire’ cannot be simply deconstructed using combinatorial means and must be considered a lexical item. To a large extent, the syntactic complexity of texts by expert and non-expert adults is due to the frequent occurrence of such lexico-syntactic or prefabricated units (Shlesinger & Ravid 2003). We will argue in this paper that the very nature of the mental lexicon changes in development to accommodate such constructions

The nature and role of N-N and N-A_{den}s places them at the focus of later language acquisition investigation, and point at the written modality as a promising domain of investigation. Their syntactic and semantic features and high lexicalizability makes them good candidates for labeling complex entities, lexemes typical of literate,

written discourse in specific text types. This brings us to the current study. Previous studies show that N-N compounding is a common device in Hebrew, cutting across genres, modalities, and registers. A_{den} s, in contrast, are a marked, literate device sometimes fulfilling similar semantic functions as N-N compounding, yet typical of densely informative, objective, hierarchically complex expository writing, on the one hand, and of older and literate writers, on the other hand. Thus, our first question was how N-N compounding and A_{den} formation patterned in non-expert Hebrew writing, and how this patterning changed in development. Secondly, we were interested in mapping the same distribution in encyclopedic texts, which are supposed to be adjusted for the reading level of their target reading population. Our second question was thus to what extent the patterning of these study constructions changed in school-related texts, indicating coordination with the morpho-syntactic development of their users.

2.0 The study

The current study compared the distribution of N-N compounds and $-i$ suffixed denominal adjectives and their sub-types in written texts produced by non-expert writers - children, adolescents and adults – with their distribution in expert-written encyclopedic texts.

2.1 Population

The study population consisted of 90 participants in 3 age groups: 6th graders aged 11;5-13;1, mean age 11;11 (17 boys, 13 girls); 11th graders aged 15;11-17;6, mean age 16;6 (13 boys, 17 girls); Adults (students and college graduates) aged 22;8-49;2 (14 men, 16 women). All participants were monolingual Hebrew speakers from a middle-high socioeconomic background, living in the center of Israel.

2.2 Procedure

Non-expert writers. Participants were asked to produce two written texts. One was a biographical text describing the life of a well-known public figure or of a friend or a family member. The second text was expository in nature, and its topic was selected by the participant out of a given list of ten topics: *War, cinema, the city of Tel Aviv, clocks, cats, higher academic studies, football, the PLO, the zoo, and sculpture.* Biographies are close to narratives since they are episodic in nature and focus on events (Berman & Slobin 1994, Hickman 1995). Expository texts focus on entities, processes and ideas (Britton 1994, Mosenthal 1985). Participants were asked to write their compositions as though they were writing an encyclopedic entry, thus directing them towards using higher-register, literate, written discourse style.

Expert writers. Three encyclopedias were selected for text analysis, each explicitly targeted for use in each of the age groups under discussion: 1) *Yavne* (1992), an encyclopedia for gradeschoolers designed after the French Larousse; 2) *Britannica for Youth* (1993), an encyclopedia for highschoolers, translated and adapted for Israeli students; and 3) *The Hebrew Encyclopedia* (1988), an adult-level encyclopedia with entries written by experts in their domains. To preserve thematic unity with the non-expert texts, the encyclopedic texts selected for analysis also consisted of (1) 10 biographical entries on world-renowned figures (e.g., Ben Gurion, Freud, Saddat, da Vinci), and (2) 10 expository entries discussing the same topics enumerated above.

2.3 Materials

The 90 non-expert writers each produced 2 texts, yielding altogether 180 texts: 90 biographical texts and 90 expository texts. The three target encyclopedias yielded altogether 60 texts: 30 biographical texts and 30 expository texts. These texts written by experts and non-experts served as the basis for our analysis.

2.4 Analysis

In preparation for a statistical analysis, we classified two linguistic types of information about the texts under investigation: 1) clause length; 2) the complex nominal constructions under consideration – N-N compounds and denominal adjectives (A_{den} s).

2.4.1 Analysis of clause length

Non-expert texts produced by the study participants were of different lengths, but all encyclopedic texts consisted of 400 words each so as to permit type / token analysis (Richards & Malvern 1997). Therefore, *text size* across expert-and non-expert texts was measured by *mean clause length*, defined as number of words divided by number of clauses (words defined for Hebrew as graphemic sequences separated by spaces, and clauses as defined in Berman & Slobin 1994). This measure has been shown to be a reliable diagnostic of textual development (Berman & Ravid 1999, Berman & Verhoeven 2002).

2.4.2 Analysis of nominal constructions

This analysis was similar across all text types produced by both experts and non-experts.

N-N compounds. N-N compounds were counted in two stages.

1. All tokens and all types were counted in each text. Compounds were lemmatized, that is, inflectional differences were disregarded.
2. All N-N compounds were classified into two types: *lexicalized* and *novel* compounds. Lexicalized compounds are those rote-learned and retrieved by Hebrew speakers as idiomatic expressions or single lexemes, e.g., *beyt[^]séfer* ‘Lit. house of books = school’, as well as names of places, of books, of works of art, e.g., *drom[^]afrika* ‘South Africa’. Six judges (all

veteran teachers of Hebrew) judged compounds as either lexicalized or novel, following the procedure described in Ravid & Shlesinger (1995).

The remaining syntactically-derived compounds were designated *novel*.

Denominal adjectives. Denominal adjectives were counted in three stages.

1. All tokens and all types were counted in each text. A_{denS} were lemmatized, that is, inflectional differences were disregarded.
2. A_{denS} were classified into two syntactic types: *predicating* adjectives, serving as predicate heads, e.g., *hu haya šriri* ‘He was muscular’; and *NP modifiers*, e.g., *séret ti’udi* ‘film documentary = documentary film’.
3. A_{denS} were also classified into those with *Hebrew* and *foreign* lexical bases, e.g., *yomi* ‘daily’ from *yom* ‘day’ (Hebrew-based A_{den}), *socialisti* ‘socialist’ (foreign-based A_{den}).

10% of all expert and non-expert texts were analyzed by a separate judge (veteran Hebrew teacher) in addition to the authors. 92% reliability was achieved, and cases of discrepancy were resolved by discussion.

2.5 Predictions

For all measures, we predicted longer, more complex and novel structures in expert texts, in older age groups, and in expository rather than biographic texts.

Expertise was predicted to be significant due to a larger and more flexible mental lexicon and better access to complex morpho-syntactic operations in expert writers, and an adherence to a consistently higher register than non-expert texts (Berman & Ravid 1999, Shalom 2002, Shlesinger & Ravid 2003).

Development was predicted to be significant in *non-expert participants* due to natural language development coupled with schooling, resulting in more linguistically literate constructions. Since school-related *expert texts* serve as input to language and

literacy development in students, we hypothesized that encyclopedia writers are sensitive to the reading level of their target audiences and adjust their language to increasing linguistic and cognitive capabilities in these audiences, resulting in similar developmental changes in *expert texts*.

Genre was predicted to be significant in both populations, given that expository texts are lexically denser and syntactically more complex than narrative texts (Berman & Verhoeven 2002), focusing on conceptual content and requiring lexically-specific terminology (Ravid in press).

3.0 Results

In this section we presents the results of the statistical analyses comparing texts produced by non-experts versus experts. This section is organized as follows: We first compare clause length in texts (3.1), then proceed to compare N-N compounds across texts (3.2), A_{denS} across texts (3.3), and finally compare choice of either construction as expressors of complex sub-categorization (3.4).

3.1 Comparison of clause length

Table 1 presents a comparison of mean clause length (number of words divided by number of clauses per text) in expert and non-expert texts, by age group and genre.

INSERT TABLE 1 ABOUT HERE

We predicted mean clause length to increase with age, genre and expertise. To test our predictions we conducted a three-way ANOVA: expertise (2: expert vs. non-expert texts) x age group (3: 6th graders / *Yavne*, 11th graders / *Britannica*, adults / *Hebrew Encyclopedia*) x genre (2: biographic vs. expository texts). An effect for expertise emerged ($F(1,228)=201.59, p<.01$): Mean clause length was higher in expert ($M=7.48$) than in non-expert texts ($M=5.32$). An effect for age emerged

($F(2,228)=13.31, p<.01$). This finding was mitigated by an interaction between expertise and age ($F(2,228)=11.63, p<.01$), depicted in Figure 1.

INSERT FIGURE 1 ABOUT HERE

This interaction shows that although expert clauses are always longer, the gap between 6th graders and their target encyclopedia (*Yavne*) is larger (3.18) than the gap between 11th graders and their target encyclopedia (*Britannica*) (1.52) and the gap between adults and their target encyclopedia (*Hebrew Encyclopedia*) (1.76). Finally, there was an effect for genre ($F(1,228)=15.90, p<.01$).

Looking closely at Table 1, it is not clear whether the age and genre effects are shared by both expert and non-expert texts. We therefore conducted a separate 2-way ANOVA on mean clause length in the encyclopedia texts: Target encyclopedia (3: *Yavne, Britannica, Hebrew Encyclopedia*) x genre (2: biographic vs. expository texts). This analysis did not produce any significant effects nor any interactions. Therefore, the effects found stem from the non-expert population. Thus, our predictions concerning mean clause length were only partially confirmed: Mean clause length increases in non-expert but not in expert texts; non-expert expositorys contain longer clauses than do biographic texts, but there is no such genre difference in expert texts; and expert clauses are always longer than non-expert clauses, with a greater discrepancy between student population and target encyclopedia in grade 6.

3.2 Comparison of N-N compounds across texts

N-N compounds were classified into two types (2.4.2 above): *lexicalized* and *novel* compounds. We predicted a decrease in lexicalized compounds and an increase in novel compounds with age, expertise and genre. To test this hypothesis, we examined the distribution of lexicalized compounds (which were more numerous than novel compounds) across the texts. Since expert texts were of the same length while

non-expert texts were of different lengths, we examined this distribution as the proportion of lexicalized compound tokens out of the total number of compound tokens. Table 2 lists these proportions by expertise, age and genre.

INSERT TABLE 2 ABOUT HERE

To test our predictions, a three-way ANOVA was conducted on the proportion of lexicalized compounds in the texts: expertise (2: expert vs. non-expert texts) x age group (3: 6th graders / *Yavne*, 11th graders / *Britannica*, adults / *Hebrew Encyclopedia*) x genre (2: biographic vs. expository texts). An effect for expertise emerged ($F(1,227)=12.27, p<.01$): Non-expert writers have a higher proportion of lexicalized compound tokens ($M=0.66$) than expert writers ($M=0.52$). An effect for age emerged ($F(2,227)=3.75, p<.05$). Finally, there was an effect for genre ($F(1,227)=12.27, p<.05$).

Again, an examination of Table 2 led to a separate 2-way ANOVA on the proportion of lexicalized compounds in the encyclopedia texts: Target encyclopedia (3: *Yavne*, *Britannica*, *Hebrew Encyclopedia*) x genre (2: biographic vs. expository texts). This analysis did not produce any significant effects nor any interactions. Therefore, the effects found stem from the non-expert population. Thus, our predictions concerning the distribution of lexicalized versus novel compounds were only partially confirmed. Non-expert texts contain fewer lexicalized compounds (i.e., more novel compounds) with age, while expert texts retain the same proportion across the board; non-expert expositorys contain fewer lexicalized compounds than do biographic texts, but there is no such genre difference in expert texts. And finally, expert texts contain a lower proportion of lexicalized compounds than do non-expert texts.

3.3 Comparison of A_{den} s across texts

Our analysis of A_{denS} related to two aspects: (i) their syntactic function as either predicate heads or NP modifiers; and (ii) their nominal base – either Hebrew or foreign.

3.3.1 *Syntactic function of A_{denS}* . We predicted that there would be more A_{denS} as NP modifiers with age, expertise and genre. Due to the difference in text lengths (see 3.2 above), this distribution was measured by the proportion of NP-modifying A_{den} tokens out of all A_{den} tokens. Table 3 lists these proportions by expertise, age and genre.

INSERT TABLE 3 ABOUT HERE

To test our predictions about the distribution of the syntactic functions of A_{denS} (2.5 above), we conducted a three-way ANOVA: expertise (2: expert vs. non-expert texts) x age group (3: 6th graders / *Yavne*, 11th graders / *Britannica*, adults / *Hebrew Encyclopedia*) x genre (2: biographic vs. expository texts). An effect for expertise emerged ($F(1,179)=12.28$, $p<.01$): The proportion of NP-modifying A_{den} tokens out of all A_{den} tokens was higher in expert ($M=0.98$) than in non-expert texts ($M=0.89$). An effect for age emerged ($F(2,179)=7.85$, $p<.01$). This finding was mitigated by an interaction between expertise and age ($F(2,179)=4.58$, $p<.05$), depicted in Figure 2.

INSERT FIGURE 2 ABOUT HERE

This interaction shows that the proportion of NP-modifying A_{denS} in 6th graders' texts is lower than in their target encyclopedia, while no such gap exists between these proportions in 11th graders' texts and their target encyclopedia and in adults' texts and their target encyclopedia.

We conducted a separate 2-way ANOVA on the proportion of NP-modifying A_{denS} in the encyclopedia texts: Target encyclopedia (3: *Yavne*, *Britannica*, *Hebrew Encyclopedia*) x genre (2: biographic vs. expository texts). This analysis did not produce any significant effects nor any interactions. Therefore, the effects found stem

from the non-expert population. Our predictions concerning the distribution of NP-modifying versus predicating A_{denS} were confirmed only partially: The developmental increase in the proportion of NP-modifying at the expense of predicating A_{denS} occurs only in the non-expert population. Expert texts targeting all age groups contain almost only NP-modifying A_{denS} . Expert writers targeting their encyclopedic text at 6th graders use NP-modifying A_{denS} to a larger extent than do 6th graders themselves in their writing, but counter to our prediction, no such gap exists between older age groups and their target encyclopedias.

3.3.2 *Lexical base of A_{denS}* . We predicted that there would be more foreign-based A_{denS} with age, expertise and genre. Due to the difference in text lengths (see 3.2 and 3.3.1 above), this distribution was measured by the proportion of Hebrew-based A_{den} types out of all A_{den} types (the foreign base is the corollary). Table 4 lists these proportions by expertise, age and genre.

INSERT TABLE 4 ABOUT HERE

To test our predictions about the distribution of the lexical base of A_{denS} (2.5 above), we conducted a three-way ANOVA: expertise (2: expert vs. non-expert texts) x age group (3: 6th graders / *Yavne*, 11th graders / *Britannica*, adults / *Hebrew Encyclopedia*) x genre (biographic vs. expository texts). However, our predictions were not confirmed. There were no significant effects nor any interactions. The Hebrew base is the most dominant in A_{denS} across the board in all age groups, text types and expertise levels.

3.4 Comparison of N-N compounds and N- A_{denS} as primary expressors of complex sub-categorization

Our final analysis concerns the status of N-N compounds and N- A_{denS} as structural devices expressing complex sub-categorization. The domain of analysis

here was the total of all novel N-N compound types plus all A_{den} types serving as NP modifiers, together making up the total of complex NP constructions expressing sub-categorization. Our prediction was that compounds would prevail overall, but that there would be a higher proportion of N- A_{den} s out of the total of complex sub-categorizing NP constructions with age, expertise and genre (N-N compounds are the corollary). Table 5 lists these proportions by expertise, age and genre.

INSERT TABLE 5 ABOUT HERE

It is clear from Table 5 that N-N compounding is the leading structure in the expression of sub-categorization in NP constructions across age groups and text-types. To test our predictions about the changing pattern of N- A_{den} distribution within the total of complex sub-categorizing NP constructions (2.5 above), we conducted a three-way ANOVA: expertise (2: expert vs. non-expert texts) x age group (3: 6th graders / *Yavne*, 11th graders / *Britannica*, adults / *Hebrew Encyclopedia*) x genre (2: biographic vs. expository texts). There was no effect for expertise: Encyclopedia writers did not use a higher proportion of N- A_{den} s than novel N-N compounds as expressors of complex sub-categorization than non-expert writers. There emerged an effect for age ($F(2,225)=14.70$, $p<.01$): In 6th grader / *Yavne* texts there was a lower proportion of N- A_{den} s out of all complex sub-categorizing NP constructions ($M=0.13$) than in the two older groups (11th graders / *Britannica* $M=0.23$, adults / *Hebrew Encyclopedia* $M=0.27$). And there was also an effect for genre ($F(1,225)=4.47$, $p<.05$).

A separate 2-way ANOVA on the proportion of NP-modifying A_{den} s in the encyclopedia texts: Target encyclopedia (3: *Yavne*, *Britannica*, *Hebrew Encyclopedia*) x genre (2: biographic vs. expository texts) did not yield any significant effects nor any interactions. Therefore, the effects found stem from the non-expert population.

Thus our predictions concerning the distribution of NP-modifying A_{den} s versus novel N-N compounds as expressors of complex sub-categorization were only partially confirmed: There are developmental changes only in the non-expert population, with an increase in the proportion of NP-modifying A_{den} s at the expense of N-N compounds from gradeschool to adults, but not in the target encyclopedias.

Expository texts contain relatively more NP-modifying A_{den} s than biographic texts, but then again, only in the non-expert texts. And finally, expert and non-expert texts do not differ in this respect.

4.0 Discussion

This study examined the distribution of complex nominal constructions in the writing of children, adolescents and adults compared with their distribution in encyclopedic texts targeting these age groups. Participants wrote biographic and expository texts, which were then compared with encyclopedic texts sharing the same thematic content. Our analysis focused on the distribution of two types of complex nominal constructions in these texts: N-N compounds (e.g., *maxane^cava* ‘camp^army = army camp’) and *i*-suffixed denominal adjectives (e.g., *maxane cva'i* ‘military camp’). Our concern was both with the development of these constructions are viewed in non-expert writing, as well as with to the extent to which these developmental changes matched changes in the distribution of these constructions in expert-written texts targeting particular age groups.

4.1 Expertise, development, and genre in textual production

We first discuss those dependent variables of our study that were analyzed in the results section.

4.1.1 Clause length in textual development and in expert texts

Since non-expert texts were of different lengths within and across age groups, we measured development in text size by mean clause length. Longer clauses consist of more words per clause, indicating both lexical and syntactic development. Lexical development is indicated since literate, advanced vocabulary denoting complex temporal, logical and causal configurations is multi-lexemic, e.g., Hebrew *ela im ken* 'unless' which would be counted in our study as three words. Mean clause length is also related to syntactic complexity: A longer clause may contain a larger number of optional phrases, in most cases intra-sentence modifiers such as adjectival and adverbial phrases, and / or longer and more internally complex phrases such as heavy NPs. Mean clause length is thus associated with lexical and syntactic growth, which are in their turn necessary for the expression and organization of information in text development.

Recall that we are interested in both changes in the distribution of our target constructions in the non-expert population, as well as in how these changes are reflected and are matched by changes in expert texts. Analysis of clause length showed that this development takes place in non-expert texts clauses, and it is related to genre: Clauses are longer in expository than in biographic texts. This means that non-expert writing becomes more elaborate and complex with age, and that students are sensitive to genre distinctions from early on. However, encyclopedia writers do not adjust clause length to their target populations, and clause length remains the same across the board in all expert texts. This means that the basic linguistic patterns in encyclopedia entries are essentially the same. Since clauses are always longer in expert than in non-expert texts, this is probably a problem for readers, and particularly for younger readers: There is a particularly large gap between clause length in 6th graders and their target encyclopedia, *Yavne*. We interpret this finding to mean that

encyclopedia writing is always more densely structured and in higher register than non-expert writing, serving as a source and an example of desired written language style (Ravid & Tolchinsky 2002). This discrepancy may be easier to handle in highschool and at university level, whereas gradeschoolers may find this discrepancy between their own syntactic construction and target written-language construction more difficult to handle.

4.1.2 N-N compounds in textual development and in expert texts

One of the study foci was the later language development of N-N compounds as viewed through children's and adolescents' writing. In order to observe developmental patterns in the current study, we examined the proportion of lexicalized compounds (the corollary of novel compounds) in the study texts. Lexicalized compounds (such as *tahalix[^]šalom* 'peace process', *oved[^]bank* 'bank worker', *tna[^]ey[^]kabala* 'acceptance requirements') comprised the majority of compounds in all age groups and expert texts, except for the Hebrew Encyclopedia (targeting educated adult Hebrew speakers). This in itself is not surprising: Lexicalized compounds are complex lexical entries and constitute an important part of the literate lexicon, complex labels for complex entities. However, in the current study we were specifically interested in those N-N compounds that were novel rather than lexicalized, i.e., did not serve as rote-learned lexical items. These were, for example, *txunot[^]ófi* 'character features', *toldot[^]ha-universita[^]ot* 'the history of universities', *nifla[^]ot[^]ha-pisul* 'the wonders of sculpture', and *švirat[^]šigra* 'breaking the routine'. An increase in the proportion of novel compounds at the expense of lexicalized ones indicates a growing ability to work the lexico-syntactic interface actively and flexibly, putting together two nouns and appropriately labeling a novel concept which comes to mind by a novel complex nominal while writing.

The pattern observed above for clause length is repeated here. There is a developmental pattern in the non-expert, but not in the expert population. Among the non-expert writers, the higher the age of the participants, the lower the proportion of lexicalized N-N compounds in their texts, that is, more novel compounds emerge. This change is particularly salient in the shift from adolescent to adult texts. However, encyclopedia texts targeting both younger and older populations contain the same proportion of lexicalized compounds. Non-expert texts are genre-sensitive: Expository texts are richer in novel compounds than biographic texts, but expert writers use the same proportion of novel compounds in their biographic and expository texts. And finally, expert texts consistently retain their edge over non-expert texts in containing fewer lexicalized compounds.

Our conclusion is that this constitutes evidence for a gradual internal change in the mental lexicons of Hebrew-speaking adolescents. Not only are they able to retrieve a large number of lexicalized compounds labeling complex entities during text writing (Anglin 1993), highschoolers' mental lexicons become increasingly flexible and maneuverable, interacting with active syntactic knowledge to produce new lexemes. These results indicate that adolescents (unlike gradeschoolers) are able to make efficient use of the stable, non-ephemeral nature of the written text to recruit both lexical and syntactic resources in the service of complex sub-categorization. Expert texts do not evidence this kind of internal growth, since professional writers possess both a rich and varied lexicon and the ability to use it flexibly in writing.

4.1.3 A_{den}s in textual development and in expert texts

Our analysis of denominal adjectives in the study texts focused on two aspects: their syntactic function (predicating vs. noun-modifying) and their lexical base (Hebrew vs. foreign).

Syntactic function. The very use of A_{denS} in any syntactic function in a text already indicates access to a literate, advanced vocabulary repository. In the current comparative analysis, we were concerned with the syntactic function of A_{denS} . As elaborated above, adjectives may function either as predicate heads or as NP-modifiers. In the first case, the A_{den} participates in creating a proposition and occupies a mandatory position. In the second case the A_{den} modifies a noun by attributing a denominal feature to it, forming a complex nominal which is then embedded in another proposition. The formation of such NP-modifying A_{denS} makes heavier demands on the cognitive and syntactic capabilities of writers. The examples of NP-modifying A_{denS} occurring in the study texts show that both nouns and adjectives participating in the construction belong to a literate, advanced vocabulary: *motxan psixologi* ‘psychological thriller’, *hon enoši* ‘human resources’, *beyt^gidul tiv’i* ‘natural habitat’, *xófeš akadémi* ‘academic freedom’.

This measure was found to be particularly diagnostic of age and schooling, independent of genre, and extremely discriminating between expert and non-expert writing. Expert texts contain virtually only NP-modifying A_{denS} , and this proportion is nearly matched by both highschool and adult texts. It is only the gradeschoolers who still make use of the predicate position, and especially (though not statistically significantly) in the biographic text. Thus, 6th graders still had to bridge the discrepancy between the lower proportion of NP-modifying A_{denS} in their own texts versus their target gradeschool encyclopedia. Our conclusion, given these facts, is that the major and predominant function of denominal *i*-suffixed adjectives in Hebrew is complex sub-categorization. While this is a very late linguistic acquisition, taking place in highschool, once this role of A_{denS} is recognized by Hebrew users,

expertise is no longer a distinguishing factor. It is a stylistic feature of literate adult Hebrew writing.

Lexical base. In many cases, Hebrew A_{denS} used in professional and academic domains take abstract foreign bases which are not Hebrew nouns, e.g., *akadémi* ‘academic’, *absoluti* ‘absolute’, *intensivi* ‘intensive’, *arxa’i* ‘archaic’, *profesyonali* ‘professional’, etc (Ravid & Shlesinger 1987, Shlesinger & Ravid 2003). Therefore we predicted an increase of foreign-based A_{denS} with age, expertise and genre. This prediction, however, was not confirmed: The majority of A_{denS} in both expert and non-expert texts of both genres are Hebrew-based, and this trend does not change significantly with age, although eyeballing non-expert texts shows a declining trend (0.87, 0.72 and 0.58 respectively). Our conclusion is that forming denominal adjectives is a mainstream *Hebrew* morphological operation, taking Hebrew nouns as bases. In order to get a closer view of foreign-based A_{denS} , specific genres such as scientific writing would have to be analyzed. An interesting note here is the close correspondence between populations and their target encyclopedias: It may be the case that writers of encyclopedia entries are implicitly sensitive to the vocabulary levels of their readers, and do not overuse lexical items beyond this level. This, however, is unlikely given the other findings of this study. Other possible reasons may be a normativistic adherence to a mainly-Hebrew lexical base, or a statistical feature of texts with literate, though non-specific vocabulary.

4.1.4 Expressing complex sub-categorization in textual development

Our results show that the predominant NP construction expressing is the novel N-N compound rather than the N- A_{den} construction, and this is true across all age groups, text types, and expertise levels. N-N compounds are acquired about age four, while denominal adjectives first appear as predicative adjectives in Hebrew child

language around age seven. NP-Modifying N-A_{den}s become a productive linguistic device only in highschool (Ravid & Zilberbuch 2003) and constitute a viable, though not primary, sub-categorization device in Hebrew. Confirming our predictions, non-expert writers indeed perceive this device as characteristic of expository writing requiring precise and elaborate description of sub-categories. These developmental changes are not shared by expert texts: NP-Modifying N-A_{den}s occupy about ¼ of all complex nominals with no genre distinction.

4.2 Input, intake and output

The story told in this study is surprising to some extent. On the one hand, we have found gradual development in a number of elements participating in written Hebrew text construction – clause length, the internal distribution of complex nominals, N-N novelty, syntactic function of A_{den}s, and their lexical base. The older and more experienced the writers, the longer their clauses, and the more complex and diverse their nominals. These distinctions are genre-sensitive even in 6th grade writing: A_{den}s, and especially NP-modifying A_{den}s, a marked, literate, later acquisition, with a lower frequency in Hebrew texts, are more typical of expository texts dealing with abstract concepts and ideas, structured around a causal framework of claims, elaboration and argumentation (Britton 1994, Mosenthal 1985). Writers' textual output thus undergoes marked developmental changes during later childhood and adolescence.

On the other hand, we have found no developmental resonance in expert texts, which constitute part of the input that novice and non-expert writers are exposed to. Encyclopedia texts, in many cases the informational basis where facts non-experts use in their compositions are found, are understandably more complex than non-expert texts. But the input window on complex nominals that we have observed in expert

texts does not match the output window. From the point of view of those measures we have investigated, all three encyclopedias are constructed alike: They all have fairly long clauses, they share the same proportion of novel compounds and the same proportion of NP-modifying A_{denS} , mostly based on Hebrew nouns. Their primary sub-categorization device is the N-N compound, and NP-modifying A_{denS} occupy a steady proportion in all encyclopedias. Thus, the school-related texts that we have examined are all written at the same linguistic level and posit a consistently difficult challenge to students of all ages. Since input and output texts are different, we may conclude that learners' intake and subsequent developmental changes in output is paced by other factors (Wijnen 2000) such as internal cognitive and linguistic pacing, distributional features of the study constructions that we have not analyzed in this work, as well as factors of saliency and transparency.

4.3 The development of nominals in the mental lexicon

The results of this study, considered in the framework of other studies, throw light on the nature and the development of nominals in the mental lexicon. Nouns in the early lexicon often denote sortals – expressions which identify and differentiate individual objects (Burke 1994, Carey & Xu 1999, Clarke in press), often linguistically encoded in single, repeated nouns. With age and schooling, the nominal lexicon increases in both type and token counts. Semantically, nouns encode less individual, more categorial and more abstract referents, take on metaphorical meanings, and become increasingly complex morphologically (Seroussi, 2003; Ravid 2003), often occurring in frozen nominal expressions or in verb-noun constructions (Ravid & Avidor 1998, Ravid & Cahana-Amitay in press). At the same time, NP structure expands and diversifies to include all a variety of compound constructions, as well as modification by A_{denS} and by relative clauses (Ravid & Shlesinger 1995,

Ravid et al. 2002, Shlesinger & Ravid, 2003). By adulthood, lexicon, morphology and syntax are intertwined. Nominals make up a large part of adult Hebrew texts. Like all lexical items in the mental lexicon of adults, they are in most cases multi-morphemic, multi-lexemic entities, standing for a cluster of related semantic notions, occurring in all syntactic functions, including “advanced” positions such as adverbials, and almost always modified in diverse and hierarchically complex ways. This is the result of amassing a large amount of real-world and school-related knowledge where complex concepts need to be expressed in complex language. In fact, nominals make up large chunks of adults’ texts, blurring the boundaries of rote-learned lexical items and syntactic constructions. We regard this as an example of *conceptualist* semantics (Jackendoff 2002), where meaning expressed by language is connected to the world as it is conceptualized by the adult individual, allowing the mental structures of real-world knowledge to interact with the mental structures of linguistic knowledge.

4.4 Educational implications

To the best of our knowledge, this is the first work of this kind, comparing linguistic constructs in students’ written output with the same constructs in school-related texts that may constitute some of the textual input that they have encountered. This study reveals that Hebrew encyclopedias are not adapted to the linguistic level of their users. Texts for younger students are as complex and contain the same proportion of complex nominals as texts for older students and for adults. No adjustment is made for younger students’ possibly lesser ability to process syntactic constructions in texts. In fact, this study has found that Hebrew schoolage-targeted encyclopedias conform to a stable, uniform linguistic mold, indifferent to age and schooling level, which involves academic language with high-register complex nominals and elaborated clause structure. This academic Hebrew style is probably one

Nominals in expert and non-expert writing

of the sources of late-emerging linguistic constructions typical of the literate Hebrew targeted by Hebrew teachers and educators. It may constitute, however, a barrier to reading comprehension in gradeschool, a topic which remains to be examined.

References

- Anderson, S.R. 1985. "Typological distinctions in word formation". In T. Shopen (ed), *Language typology and syntactic description*, Vol. III *Grammatical categories and the lexicon*. Cambridge: Cambridge University Press. 3-56
- Anglin, J.M. 1993. "Vocabulary development: A morphological analysis". *Monographs of the Society for Research in Child Development*, 58, 10.
- Aronoff, M. 1980. "Contextuals". *Language* 56: 744-758.
- Berman, R.A. 1987. "A developmental route: learning about the form and use of complex nominals". *Linguistics* 25: 1057-1085.
- Berman, R.A. 1988. "Language knowledge and language use: binominal constructions in Modern Hebrew". *General Linguistics* 28: 261- 285.
- Berman, R.A. 1997. "Preschool knowledge of language: What 5-year-olds know about language structure and language use". In C. Pontecorvo (ed), *Writing development : an interdisciplinary view*. Amsterdam: John Benjamins, 61-76.
- Berman, R.A. In press a. "Genre and modality in developing discourse abilities". In C. Moder & A. Martinovic-Ziv (eds), *Discourse across languages and cultures*. (Typological studies in Language series). Amsterdam: John Benjamins.
- Berman, R.A. In press b. "Developing discourse stance in different text types and languages". *Journal of Pragmatics*.
- Berman, R.A. & I. Katzenberger. In press. Form and Function in Introducing Narrative and Expository Texts: A Developmental Perspective. *Discourse Processes*.
- Berman, R.A. & Ravid, D. 1999. "The oral / literate continuum. Final project report for the Israeli Academy of Sciences". Tel Aviv University.
- Berman, R.A. & Ravid, D. 2003. "The development of register awareness in Hebrew". Symposium in Language and Society conference, Hebrew University, 15.6.2003.

- Berman, R.A. & Slobin, D.I. 1994. *Relating events in narrative*. Hillsdale, NJ: Erlbaum.
- Berman, R.A. & Verhoeven, L. 2002. "Cross-linguistic perspectives on the development of text-production abilities". *Written Language and Literacy* 5: 1-43.
- Borer, H. 1988. "On the morphological parallelisms between compounds and constructs". *Morphology Yearbook* 1: 45-65.
- Britton, B.K. 1994. "Understanding expository text". In M.A. Gernsbacher (ed), *Handbook of psycholinguistics*. New York: Academic Press, 641-674.
- Burke, M. B. 1994. "Preserving the principle of one object to a place: a novel account of the relations among objects, sorts, sortals, and persistence conditions," *Philosophy and Phenomenological Research* 54: 591-624
- Cahana-Amitay, D. & Ravid, D. 2000. "Optional bound morphology in the development of text production". In S.C. Howell, S.A. Fish & T. Keith-Lucas (eds), *Proceedings of the 24th Annual Boston University Conference on Language Development, Vol. I*, 176-184. Somerville, MA: Cascadilla Press.
- Carey, S. & Xu, F. 1999. "Sortals and kinds: an appreciation of John Macnamara". In R. Jackendoff, P. Bloom, & K. Wynn, (Eds.). *John Macnamara: On the Border*. Cambridge, MA: MIT Press, 311-336.
- Chafe, W. 1994. *Discourse, consciousness, and time*. Chicago: The University of Chicago Press.
- Chafe, W. & Danielewicz, J. 1987. "Properties of spoken and written language". In R. Horowitz & S.J. Samuels (eds), *Comprehending oral and written language*. San Diego: Academic Press, 83-113.
- Chomsky, N. 1970. "Remarks on nominalization". In R. Jacobs and P. Rosenbaum *Readings in English Transformational Grammar*. Waltham, MA: Blaisdell. 184-221

- Choueka, Y. 1996. *Rav Milim - the Modern Hebrew dictionary*. Tel Aviv: Center for Educational Technology. [in Hebrew]
- Clark, E.V. & Berman, R.A. 1984. "Structure and use in the acquisition of word formation". *Language* 60: 542-590.
- Clark, E.V. & Clark, H.H. 1979. "When nouns surface as verbs". *Language* 55: 767-811.
- Clark, H. 1996. *Using language*. Cambridge: Cambridge University Press.
- Clarke, D.S. In press. *Sign levels: Language and its evolutionary antecedents*. Dordrecht: Kluwer.
- Halliday, M.A.K. 1988. "On the language of physical science". In M. Gahdessy (ed), *Registers of Written English: Structural Factors and Linguistic Features*. Kent, England: Pinter Publishers, 162-178.
- Ferris, C. 1993. *The meaning of syntax: A study in the adjectives of English*. Longman: London.
- Gayraud, F. 2000. Le développement de la différenciation oral/écrit vu à travers le lexique. Thèse de Doctorat, Université de Lyon.
- Hickmann, M. 1995. "Discourse organization and the development of reference to person, space, and time". In P. Fletcher & B. MacWhinney (eds), *The handbook of child language*. Oxford: Basil Blackwell, 194-218.
- Jackendoff, R. 2002. *Foundations of language*. New York: Oxford University Press.
- Karmiloff-Smith, A. 1986. "Some fundamental aspects of language development after age five". In P. Fletcher & M. Garman (eds), *Language acquisition: Studies in first language development*. Cambridge: Cambridge University Press.
- Kress, G. 1994. *Learning to write*. London: Routledge.

- Levin, I., Ravid, D., & Rapaport, S. 2001. "Morphology and spelling among Hebrew-speaking children: From kindergarten to first grade". *Journal of Child Language* 28: 741-769.
- Lieber, R. 1983. "Argument linking and compounds in English". *Linguistic Inquiry* 14:251-285.
- Mosenthal, P.B. 1985. "Defining the expository discourse continuum". *Poetics* 14: 387-414.
- Nippold, M.A. 1998. *Later language development: the school-age and adolescent years*. Austin, Texas: Pro-Ed.
- Olson, D.R. 1994. *The world on paper*. Cambridge: Cambridge University Press.
- Ong, W. J. 1992. "Writing is a technology that restructures thought". In P. Downing, S. D. Lima & M. Noonan (eds), *The linguistics of literacy*. Amsterdam: John Benjamins, 293-319.
- Paltridge, B. 1997. *Genre, frames and writing in research setting*. Amsterdam: John Benjamins.
- Pontecorvo, C. 1997. "Studying writing and writing acquisition today: A multidisciplinary view". In C. Pontecorvo (ed), *Writing development: An interdisciplinary view*. Amsterdam: John Benjamins, xv-xxxii.
- Ravid, D. 1997. "Between syntax and the lexicon: the parallel between N-N compounds and N-A strings in acquisition". In A. Sorace, C. Heycock & R. Shillcock (eds.) *Proceedings of the GALA '97 conference on language acquisition*. Edinburgh: University of Edinburgh, 138-141.
- Ravid, D. 2003. "The Noun scale in text-related development". Paper presented at Georgetown University Round Table on Language and Linguistics. Georgetown University, Washington, February 15-17.

- Ravid, D. In press. "Emergence of linguistic complexity in written expository texts: Evidence from later language acquisition". In D. Ravid & H. Bat-Zeev Shyldkrot (eds.) *Perspectives on language and language development*. Dordrecht: Kluwer.
- Ravid, D. & Avidor, A. 1998. "Acquisition of derived nominals in Hebrew: developmental and linguistic principles". *Journal of Child Language* 25: 229-266.
- Ravid, D. & D. Cahana-Amitay. In press. "Verbal and nominal expression in narrating conflict situations in Hebrew". *Journal of Pragmatics*.
- Ravid, D. & M. Nir. 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*. Groningen: Center for Language and Cognition, 113-124.
- Ravid, D. & Y. Shlesinger. 1987. "Deriving denominal adjectives in Modern Hebrew". *Hebrew Linguistics* 25: 59-70. [in Hebrew]
- Ravid, D. & Y. Shlesinger. 1995. "Factors in the selection of compound-type in spoken and written Hebrew". *Language Sciences* 17: 147-179.
- Ravid, D. & S. Zilberbuch. 2003. "Morpho-syntactic constructs in the development of spoken and written text production". *Journal of Child Language* 30: 1-24,
- Ravid, D. & Tolchinsky, D. 2002. "Developing linguistic literacy: A comprehensive model". *Journal of Child Language* 29: 419-448.
- Ravid, D., van Hell, J., Rosado, E., & Zamora, A. 2002. "Subject NP patterning in the development of text production: A crosslinguistic study of Dutch, English, Hebrew, and Spanish". *Written Language and Literacy* 5: 69-94.
- Richards, B. J. & Malvern, D.D. 1997. *The new Bulmershe Papers. Quantifying lexical diversity in the study of language development*. Reading: The University of Reading.

- Schneuwly, B. 1997. "Textual organizers and text types: ontogenetic aspects in writing". In J. Costermans & M. Fayol (eds), *Processing interclausal relationships: studies in the production and comprehension of text*. Mahwah, NJ: Erlbaum, 245-263.
- Scott, C.M. 1988. "Spoken and written syntax". In M.A. Nippold (ed), *Later language development: Ages nine through nineteen*. Austin, TX: Pro-Ed, 49-95.
- Seroussi, B. 2003. Derived nominals in the mental lexicon in a developmental perspective. MA thesis, Department of Communications Disorders, Tel Aviv University.
- Shalom, T. 2002. Linguistic structures and discourse stance in Hebrew L₁ textbooks. Doctoral dissertation, Bar Ilan University. [in Hebrew]
- Shlesinger, Y. & S. Ravid. 2003. "The distribution of adjectives in expert and non-expert texts". Paper presented at the AAAL meeting, Arlington University, Virginia, March 22-25.
- Slobin, D.I. 1977. "Language change in childhood and in history". In J. MacNamara (ed), *Language Learning and Thought*. New York: Academic Press. 185-214
- Smith, M.L. 1998. Sense and sensitivity: An investigation into fifth grade children's knowledge of English derivational morphology and its relationship to vocabulary and reading ability. Unpublished doctoral dissertation, Graduate School of Education, Harvard University.
- Strömquist, S. & Wengelin, A. 1999. "The temporal patterning of actions in on-line writing". In R. A. Aisenman (ed.) *International Literacy Project Working Papers in Relevant Features of Spoken and Written Language. Developing Literacy across Genres, Modalities, and Languages, Vol. I*. Tel Aviv: Tel Aviv University Press, 248-256.
- Wijnen, F. 2000. "Input, intake and sequence in syntactic development". In M. Beers, B. van den Bogaerde, G. Bol, J. de Jong, & C. Rooijmans (eds), *From sound to sentence:*

Nominals in expert and non-expert writing⁴⁰

Studies on first language acquisition. Groningen: Center for Language and Cognition,
164-186.

Expertise	Age Group	Genre			
		Biographic		Expository	
		Mean	SD	Mean	SD
Experts (Encyclopedias)	Yavne	7.56	1.09	7.98	1.37
	Britannica	7.22	0.96	7.11	0.80
	Hebrew Enc.	7.11	0.79	7.87	0.94
Non-experts (Participants)	6 th Graders	4.38	0.76	4.81	0.95
	11 th Graders	5.36	1.00	5.92	0.97
	Adults	5.36	1.05	6.10	1.30

Table 1. Means and standard deviations of mean clause lengths (number of words divided by number of clauses) per text, by expertise (encyclopaedic vs. participants' texts) by age group (6th graders, 11th graders, adults; *Yavne*, *Britannica*, *Hebrew Encyclopedia*), and by genre (biographic vs. expository texts).

Expertise	Age Group	Genre			
		Biographic		Expository	
		Mean	SD	Mean	SD
Experts (Encyclopedias)	Yavne	0.60	0.14	0.51	0.13
	Britannica	0.53	0.16	0.52	0.11
	Hebrew Enc.	0.51	0.12	0.42	0.08
Non-experts (Participants)	6 th Graders	0.73	0.26	0.69	0.36
	11 th Graders	0.67	0.24	0.67	0.27
	Adults	0.68	0.16	0.54	0.25

Table 2. Means and standard deviations of the proportion of lexicalized compound tokens out of the total number of compound tokens per text, by expertise (encyclopaedic vs. participants' texts) by age group (6th graders, 11th graders, adults; *Yavne*, *Britannica*, *Hebrew Encyclopedia*), and by genre (biographic vs. expository texts).

Expertise	Age Group	Genre			
		Biographic		Expository	
		Mean	SD	Mean	SD
Experts (Encyclopedias)	Yavne	0.97	0.63	0.96	0.05
	Britannica	0.98	0.03	1.00	0.00
	Hebrew Enc.	0.98	0.03	0.99	0.03
Non-experts (Participants)	6 th Graders	0.53	0.51	0.84	0.36
	11 th Graders	0.92	0.22	0.92	0.15
	Adults	0.95	0.15	0.93	0.15

Table 3. Means and standard deviations of the proportion of NP-modifying A_{den} tokens out of the total number of A_{den} tokens per text, by expertise (encyclopaedic vs. participants' texts) by age group (6th graders, 11th graders, adults; *Yavne*, *Britannica*, *Hebrew Encyclopedia*), and by genre (biographic vs. expository texts).

Expertise	Age Group	Genre			
		Biographic		Expository	
		Mean	SD	Mean	SD
Experts (Encyclopedias)	Yavne	0.71	0.12	0.63	0.15
	Britannica	0.70	0.13	0.69	0.23
	Hebrew Enc.	0.69	0.22	0.70	0.18
Non-experts (Participants)	6 th Graders	0.83	0.35	0.87	0.49
	11 th Graders	0.75	0.27	0.72	0.30
	Adults	0.75	0.34	0.58	0.33

Table 4. Means and standard deviations of the proportion of Hebrew-based A_{den} types out of the total number of A_{den} types per text, by expertise (encyclopaedic vs. participants' texts) by age group (6th graders, 11th graders, adults; *Yavne*, *Britannica*, *Hebrew Encyclopedia*), and by genre (biographic vs. expository texts).

Expertise	Age Group	Genre			
		Biographic		Expository	
		Mean	SD	Mean	SD
Experts (Encyclopedias)	Yavne	0.19	0.06	0.23	0.08
	Britannica	0.23	0.05	0.21	0.06
	Hebrew Enc.	0.28	0.09	0.23	0.10
Non-experts (Participants)	6 th Graders	0.06	0.13	0.15	0.20
	11 th Graders	0.23	0.21	0.32	0.23
	Adults	0.22	0.10	0.23	0.15

Table 5. Means and standard deviations of the proportion of N-A_{den} types out of the total number of complex sub-categorizing NP constructions per text, by expertise (encyclopaedic vs. participants' texts) by age group (6th graders, 11th graders, adults; *Yavne*, *Britannica*, *Hebrew Encyclopedia*), and by genre (biographic vs. expository texts).

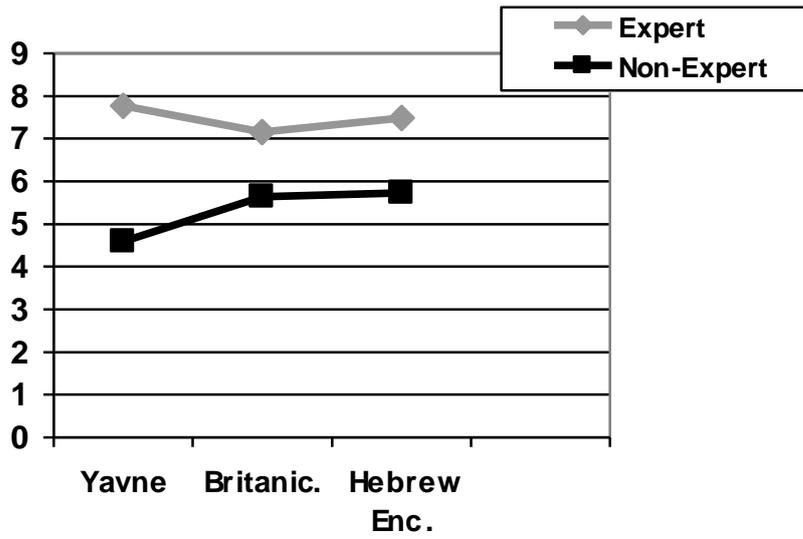


Figure 1. Interaction of mean clause length by expertise and age

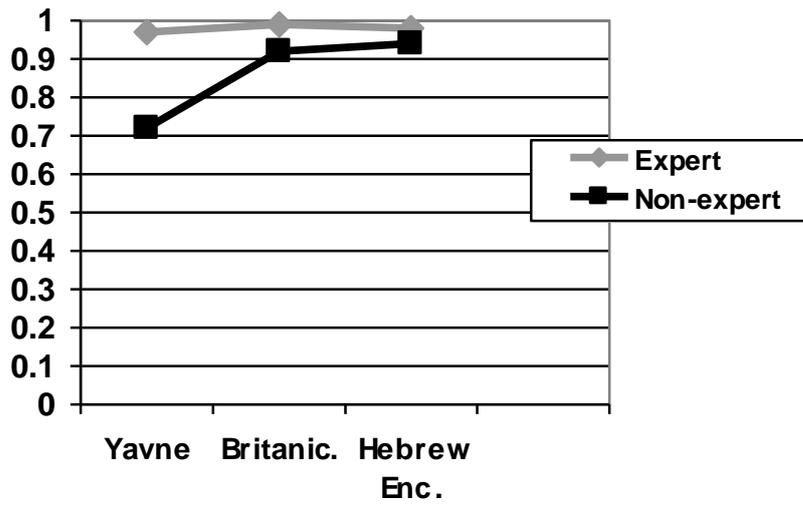


Figure 2. Interaction of NP-modifying A_{den}S by expertise and age

