First Language

Developing Noun Phrase Complexity at School Age: A Text-Embedded Cross-Linguistic Analysis

Dorit Ravid and Ruth A. Berman

The online version of this article can be found at:

Published by:

$\bigcirc$SAGE

Additional services and information for *First Language* can be found at:

Email Alerts: http://fla.sagepub.com/cgi/alerts

Subscriptions: http://fla.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.co.uk/journalsPermissions.nav

Citations http://fla.sagepub.com/cgi/content/refs/30/1/3

Developing Noun Phrase Complexity at School Age: A Text-Embedded Cross-Linguistic Analysis

Dorit Ravid and Ruth A. Berman

Tel Aviv University

Abstract

Development of noun phrase structure and use is analyzed as an important facet of syntactic acquisition from middle childhood to adolescence. Noun phrases occurring in narrative and expository texts produced in both speech and writing by 96 native speakers of English and Hebrew were identified and examined by a set of specially devised criteria including length in words, syntactic depth, abstractness of head nouns, and nature of modifiers. Results reveal a clear and consistent developmental increment in NP complexity from age 9 to 12, and particularly from age 16 years; written expository texts emerge as a favored site for use of syntactically complex constructions; and nominal elements play a more central role in the discursive syntax of Hebrew than English. Findings are discussed in terms of the interplay between psycholinguistic factors of cognitive processing constraints and the impact of increased literacy in later language development.

Introduction

The domain of 'complex syntax' is typically associated with the traditional notion of complex sentences (Jespersen, 1909; Lyons, 1977), in linguistics (Bybee & Noonan, 2002; Cristofaro, 2003; van Valin, 2006), in language acquisition (Diessel, 2004; Lust, Foley, & Dye, 2008), and in pedagogically and clinically motivated research (Hunt,

Corresponding author:
Dorit Ravid, School of Education, Tel Aviv University, Tel Aviv 69978, Israel; email: doritr@post.tau.ac.il

1965; Scott, 2004). These studies share a focus on clause combining and subordination as critical facets of later-developing syntactic abilities. The research reported in this article examines the development of complex syntax from a rather different perspective, by analyzing noun phrases (NPs) as a platform for constructing broader, discourse-embedded syntactic architecture.

Syntactic development in early language acquisition has been studied mainly in naturalistic language samples, for example, in English by Brown (1973), Radford (1990), and Tomasello (2003). Psycholinguistic analyses of later syntax, starting with Carol Chomsky's (1969) pioneering study, generally rely on experimental paradigms (Ravid & Saban, 2008), most typically in comprehension (e.g., Friedmann & Novogrodsky, 2004). The present study examines text-embedded syntactic production, in the conviction that the best way to evaluate later-developing linguistic abilities is in the context of extended discourse (Berman, 2008; Ravid & Zilberbuch, 2003a, 2003b). Research on the text construction abilities of adolescents suggests that no experiment can capture the full richness and diversity of syntactic constructions speaker-writers recruit to meet the needs of different communicative settings (Berman, 2005; Berman & Nir-Sagiv, 2008; Jisa, 2004; Ravid, 2004b; Tolchinsky & Rosado, 2005). The study presented here examines development of NP structure and use as shedding light on syntactic acquisition from middle childhood to adolescence in English and Hebrew, two typologically distinct languages.

In language acquisition, NPs manifest a gradual increase in length and complexity over time, from the typically one-word, pragmatically motivated constituents of early child utterances (*doggy*, *milk*) to preschoolers' syntactically well-formed structures (*this dog*, *my milk*). That is, the grammar of NPs in different languages is basically in place by 3 years of age (Radford, 1990; Slobin, 1985). Our concern here is with elaborations that take place subsequently, during the school years, in both the NP head and its associated modifiers, as reflecting three related developments: in lexical repertoire, syntactic proficiency, and communicative competence. To the best of our knowledge, the role of noun phrases in these developments has been the topic of little research to date.

A related goal of the study is to demonstrate how the NP can serve as a diagnostic tool for examining the psycholinguistics of later syntactic development. As constructions, NPs are mid-level in size, smaller than a clause or sentence, but easily extendable beyond a single word. Moreover, they are involved in both clause-internal structure (as subject or as object of a verb or preposition) and inter-clause syntactic linkage (in use of relative clauses). As units of linguistic processing, NPs have clearly defined syntactic boundaries and typically fixed ordering of internal elements, not accessible to syntactic extraction. In consequence, they impose relatively little burden on short-term memory and hence on learning, as reflected in the fact that they are well formed from early on. Even 3-year-olds preserve the internal ordering of elements in the NP, and can construct them correctly, and this is certainly true of grade-school children, the youngest participants in the study reported in this article. In fact, in oral texts produced by Hebrew-speaking 9- to 10-year-olds, NP structures were among the few areas of syntactic structure where their language output was fully grammatical and not at all chaotic (Ravid, 2006). Besides, NPs as units of analysis are readily comparable across speech and writing, in contrast to more conventional linguistic constructs like *word* (consider contractions in English or liaisons in French) or *sentence*, which is notoriously difficult to identify in speech production.

Particularly relevant to the current undertaking is the fact that NPs as syntactic phrases consist of an obligatory head plus optional modifiers (Jackendoff, 1977). This means that they can be more or less complex both in the nature of the head noun and in the number and type of its associated modifiers – a feature that has clear developmental consequences. With age and schooling, the nouns that serve as NP heads are more abstract and categorial in nature, while modifiers increase in both amount and the types of dependencies they express. For example, heavy NPs often contain prepositional phrases (PPs), yielding nominal strings of the kind that no young child would produce (e.g., *the bulk of the information in our textbook*). This observation accords with the insight of school-oriented evaluations suggesting that general text quality can be tapped by counting the number of PPs (Loban, 1976).

The current study considers the development of NP complexity when anchored in extended discourse. Prior research on school-age monologic texts in the cross-linguistic project mentioned earlier has highlighted the role of *text type* in determining various facets of linguistic expression – morphology, lexicon, register, syntactic connectivity, as well as discourse structure and content (Berman, 2008; Ravid & Berman, 2009; Tolchinsky & Rosado, 2005). Here, 'text type' refers to the intersection of two dimensions of discourse – genre and modality. Genre relates to social-cultural practices (Grimshaw, 2003; Paltridge, 1997); it is determined by communicative setting, and reflects different cognitive representations (Bruner, 1986; Steen, 1999). The two types of discourse analyzed later in this article – narrative and exposition – represent two clearly distinct genres. Narratives focus on people, their actions and motivations, and they express the unfolding of events in a temporal framework (Berman & Slobin, 1994; Longacre, 1996), whereas expository texts focus on topics and issues, and they express the unfolding of ideas, claims, and arguments with logical interrelations between them (Boscolo, 1990; Britton, 1994; Mosenthal, 1985). The two genres analyzed later thus reflect the division proposed by Bruner (1986) between narrative and logico-scientific discourse as embodying two distinct modes of thought.

We have found, further, that the effect of discourse genre is markedly mediated by the factor of modality, that is, of spoken versus written language (Berman & Ravid, 2009). These two modalities represent different modes of language production, since speech is produced under pressure of rapid online processing, whereas the displaced time and offline production of writing allows planning and monitoring (Chafe, 1994; Clark, 1996; Strömqvist, Nordqvist, & Wengelin, 2004). The two modalities also represent distinct styles of discourse, since spoken language is typically more everyday, colloquial, personalized, and interactive, while written language represents a depersonalized 'discourse stance' (Berman, 2005), one that is more formal and detached (Biber, 1998; Ravid & Tolchinsky, 2002) and better planned and more organized (Berman & Nir-Sagiv, 2007; Ravid & Berman, 2006). Consequently, written language is typically better organized structurally, grammatically more complex, and richer in lexical density and variety (Halliday, 1989; Olson, 2006), even in stories produced by young school-age children (Cameron, Hunt, & Linton, 1988).

Target language typology is an additional factor that may impinge on NP structure and use. English and Hebrew are two languages that we assume will reveal interesting text-embedded differences in this respect. For example, English manifests elaborate verb phrase structures

with auxiliary and other helping verbs, along with a rich array of verb plus satellite 'phrasal verb' constructions. In contrast, Hebrew is a verb-framed language (Slobin, 2004) with morphologically complex head verbs, but whose verb phrases are syntactically sparse (Berman, 1980a); Hebrew abounds in verbless so-called 'nominal sentence' structures and, in the absence of a lexically distinct class of manner adverbs, relies heavily on nominally based adverbials (Berman, 1978). Compare, for example, Hebrew *bi-zehirut raba* 'with great carefulness' with English *very carefully* (Berman, 1999; Ravid & Shlesinger, 1999). This suggests that NP complexity may feature more prominently in Hebrew than in English.

Noun Phrase Complexity

Against this background, we analyzed NPs as a yardstick of syntactic complexity across the variables of development, genre, modality, and cross-linguistic variation. The account of 'NP complexity' presented in this section is motivated both qualitatively and quantifiably by the psycholinguistic view of grammar and lexicon as interrelated facets of language knowledge and use (Bates & Goodman, 1997). Evidence for the strong connection between command of lexicon and syntax has been shown for young preschoolers (e.g., Marchman & Bates, 1994; Marchman & Thal, 2005) as well as for school-age language development (Berman & Nir-Sagiv, 2007; Ravid, 2004a; Ravid & Cahana-Amitay, 2005). Our working hypothesis is thus that lexical semantics will play a role in the development of NP constructions, and that both content and structure need to be considered in this as in other linguistic domains (Goldberg, 1995; Nir-Sagiv, 2008).

In the present context, an NP is defined as a string of words with a lexical noun as its head. At a surface level, NP complexity can be identified with *length*, specified in this study both by mean number of words per NP and by the mean amount of NPs per text. Yet complexity is also a function of such factors as relative semantic complexity of the head noun; amount and variety of modifying elements associated with a given head NP – quantifiers, adjectives, prepositional phrases, relative clauses; and syntactic depth in the sense of how many N nodes are dominated by a single NP. In specifying what is meant by NP complexity, accordingly, all these features need to be taken into account together with how they interact with one another.

In the Method section we detail the criteria we designed to trace the development of text-embedded NP structure and content across adolescence in four text types and in two typologically different languages. We aim to demonstrate that (1) NPs are a viable diagnostic criterion of language use across the variables of development, genre, and modality; (2) they constitute a platform for school-age development of discursive syntax; and (3) their text-embedded distribution is governed by target language typology.

Noun phrases, as noted, can be analyzed from different points of view. Our focus here is on lexical noun phrases – consisting minimally of a single lexical noun as their head, for example, *boys*, *conflict*. Lexical NPs are typically expandable by one or more modifiers, such as: determiners – including articles (*the boys*), demonstratives (*this conflict*), quantifiers (*several boys*); adjectives or adjectival phrases (*extremely severe conflict*); prepositional phrases (*boys from the neighborhood*; *a conflict with one's peers*); and/or relative clauses (*a conflict that has lasted for centuries*). Complex NPs also cover nominal modification, such as noun compounds (*homeroom teacher*, *conflict management*)

and other kinds of genitives (*the teachers' agenda)*. Our study is thus confined to all and only NPs with a lexical noun as head.[1]

Further, concern here is with the internal structure and content of NPs rather than with the syntactic relation that they realize as arguments or adjuncts (Andrews, 2007), or with the role they play in expressing the discourse function of reference (Hickmann, 2003). We also exclude from analysis the question of *site*, in the sense of subject versus non-subject NPs. This topic has been considered from different perspectives in analyses of comparable data from the larger cross-linguistic project: for Spanish in relation to pre-ferred argument structure (Khourajaina & Tolchinsky, 2004); for French in comparing written and spoken usage by Jisa (2004); and in relation to Chafe's (1994) 'light subject constraint' in four different languages including Hebrew and English (Ravid, van Hell, Rosado, & Zamora, 2002). Issues of grammatical functions and thematic roles are of considerable interest to our general research agenda, but – given the numerous variables involved in the current investigation – our analysis here is deliberately confined to con-struction-internal facets of NPs. This means, inter alia, that the categories analyzed in this study concern features of NP structure and content that can a priori be considered equivalent across different languages, so excluding structural variations of interest to linguistic typology, such as head direction or subjectless constructions (Berman, 1980b, 1988, 1990; Ravid & Shlesinger, 1995; Ravid & Zilberbuch, 2003a).

Noun phrases so delimited are analyzed in this study in relation to the four variables of: (1) development – four levels of age and schooling, from middle childhood to young adulthood; (2) genre – personal experience narratives and expository discussions; (3) modality – speech and writing; and (4) target language – Californian English and Israeli Hebrew. The following predictions are examined. *Developmentally*, we should find lon-ger and more complex NPs as a function of age and schooling, with a major cut-off point between middle school and high school students, with more complex NPs constituting a hallmark of maturely proficient text construction. In *text type*, we expect to find a hierar-chical progression in NP complexity as a function of genre and modality, proceeding from spoken narratives at one extreme via written narratives and spoken expository texts in the middle to written expository texts at the other extreme. Our *cross-linguistic* predic-tions consider NPs as reflecting more general typological features of Hebrew compared with English: NPs should play an important role in developing and constructing complex syntactic architecture in both languages but we expect them to take up more 'text space' and to play a more central role in the discursive syntax of Hebrew compared with English.

Method

Our analyses in both languages derive from two parallel databases, collected in the framework of a large-scale cross-linguistic project in seven different languages.[2]

Procedures

Data collection and transcription in California and Israel followed identical procedures. All participants were first shown the same short video film without words depicting (unresolved) situations of conflict – moral, social, and physical – in a school setting.

They were then asked to write and tell a *story* about an incident where they had been involved in a situation of 'problems between people' and to write a *composition* and give a *talk* discussing ideas on the topic of 'problems between people.' In other words, each participant in each age group in both countries produced four texts: one oral and one written narrative, and one oral and one written expository text. Task order was balanced across the participants for genre (narrative/expository) and modality (spoken/written). For details, see Berman and Katzenberger (2004) and Berman and Verhoeven (2002).

Participants

The 96 participants in the study were monolingual, native speakers of Israeli Hebrew and of Californian English, from well-educated backgrounds and attending well-established schools. There were 48 participants in each country, 12 in each of the four age/schooling levels: middle childhood, aged 9–10 years, 4th grade; pre-adolescence, aged 12–13 years, 7th grade; adolescence, aged 16–17 years, 11th grade; and graduate school university students majoring in the sciences or humanities, aged 20–30 years.

Text Types

Each of the 96 participants, as noted, produced four texts (a spoken and written narrative, and a spoken and written exposition), yielding a database of 384 texts, all on the same socially relevant theme of interpersonal conflict (Tolchinsky, Johansson, & Zamora, 2002). The sample analyzed in this article thus differs from materials typically used in studies of (developing) discourse: it focuses on text construction rather than comprehension; the fact that the same person produced all four texts makes them uniquely comparable; and monologic rather than conversational discourse was selected as a preferable context for complex NP usage (Ravid, Olshtain, & Ze'elon, 2003).

Criteria for Noun Phrase Analysis

In analyzing NP structure and content, we applied a multidimensional set of five criteria, involving semantics, lexicon, and syntax, with the aim of specifying NP profiles along the variables of development, text type, and language. The method we designed addressed the issues of (1) *demarcation* of complex NPs that contain subordinate NPs headed by lexical nouns, since this is liable to lead to double counts; and (2) *NP complexity* as affected by typological variation, for example, in the case of differences in head direction or reliance on noun compounding (favored in Hebrew and Germanic languages compared with Romance).

   With respect to *demarcation*, NPs were marked off at their largest boundary, such that any given NP includes a *lexical* head and all the modifiers associated with it: determiners, adjectives, prepositional phrases, NPs, infinitival complements, and subordinate or non-finite (relative) clauses. For example, *the inquisition in Spain* is a single NP whose head is *inquisition.* And a string like *a detailed step-by-step instruction sheet explicating the analyses themselves* is also a single NP whose head is *sheet*. In other words, not every single lexical N is counted as the head of a given NP. This procedure meant we could keep N-slot lexicality (see note 1) and NP complexity apart, and so eliminate the problem of double counting.

A second issue was assessing *NP complexity*. Consider the following NPs occurring in the expository text written by a high school boy, ranging from four to seven words in length, with head nouns underlined: *a <u>result</u> of insecurity*, *another <u>source</u> of conflict*, *many <u>people</u> in this world*, *close <u>contact</u> with someone unknown*, *their dominant <u>status</u> amongst a group*, *the other person's <u>point</u> of view*, *<u>students</u> who are unfamiliar with another student*, *the <u>relationships</u> in which they are involved.* These NPs are 'heavy' since they are relatively long and they contain several different types of modifying expressions: adjectives – like *close*, *dominant*; post-nominal PPs – like *of insecurity*, *of conflict*, *in this world*, *with someone unknown*, *amongst a group*; and relative clauses. The challenge for us was to define the intuitive sense of nominal 'heaviness' in a quantifiable fashion going beyond mere length in words and comparable across different languages.

To meet these requirements, we devised a set of explicit criteria assigning to each NP in our sample a score on *a four-level ranking of complexity*, ranging from 1 = simple, 2 and 3 = medium to high level of complexity, and 4 = highly complex. Scoring was conducted by five Hebrew–English bilinguals, who each worked independently and then all met together with the two authors weekly over the entire 6-month period of coding to ensure inter-coder reliability. (Fuller details on coding instructions can be obtained from the authors.) Each NP in each of the four texts produced by participants was assigned a score from 1 to 4 for each of the following five criteria: *length* in words, *semantic complexity* of head noun, quality and number of *modifiers*, syntactic *depth*, and syntactic *variability*. These NP complexity criteria, as described and illustrated below from our database, are in principle applicable across different languages, although the specific values assigned need to be designated separately for any given language.

Criterion #1: Length in Words.

A score of 1 = short NPs, from one to three words (e.g., *an <u>individual</u>*); 2 = middle-length NPs, of four to six words (e.g., *the <u>condemnation</u> of non-Christian people*); 3 = higher level NPs, of seven to nine words in length; and 4 = long NPs, over 10 words (e.g., *a <u>case</u> where the difference in beliefs causes a major difference in lifestyle and even an inability to get along*). These criteria are different for Hebrew, which has much richer word-internal bound morphology. In Hebrew, moreover, several function words are orthographically attached to the next word, including the conjunctions *she* 'that' and *ve-* 'and,' the definite article *ha-* 'the,' and the prepositions 'in' (or 'at'), 'to,' 'from,' and 'like' (set off by a hyphen in the examples given below). As a result, Hebrew NPs have a slightly shorter length-range than their English counterparts, as illustrated by these examples from high school texts, with head nouns underlined): a score of 1 = one to two words – *<u>ha-xaverim</u> sheli* 'the-<u>friends</u> of-mine' = 'my friends'; score of 2 = three to five words – *gdoley ha-filosofim ve-ha-psixologim* 'the <u>greatest</u> of the philosophers and the psychologists'; score of 3 = six to seven words – *<u>ha-shoresh</u> shel kol inyan ha-pshaim ba-olam* 'the <u>root</u> of the whole issue of the crimes in the world' (where the English equivalent would be counted as having seven or more words); score of 4 = eight words or more, – *<u>maskana</u> meshutefet ha-kova'at ki ha-adam hu yecur xevrati* 'a shared conclusion which contends that man is a social being.'

Criterion #2: Semantic Complexity of Head.

This analysis derives from an original 10-point ranking of nominal abstractness devised by the first author and tested on the

Hebrew database of this study (Ravid, 2006), and subsequently extended to the corresponding English sample (Nir-Sagiv, Bar-Ilan, & Berman, 2008). The original 10-point semantic complexity scale was collapsed to a four-place ranking, as follows. The lowest rank of 1 is assigned to concrete objects and specific people, e.g., *John*, *a ball*, *flowers*, *a boy in my class*; a score of 2, to categorial nouns designating collections and location (such as *my class*, *the city*), role nouns such as *teacher*, *counselor*, *officer*), generic nouns (*things*, *people*), and temporal nouns (*minute*, *year*, *day*, *summer*); a higher score of 3 is assigned to non-abstract, high-register or rare nouns like *rival*, *cult*, abstract but relatively everyday terms like *fight*, *war*, to metaphorical extensions of concrete terms like *path to success*, *river of time*, and to generic temporal nouns like *time*; and the highest level of 4 is assigned to nouns that are low frequency, non-imageable and abstract, and typically derivationally complex, e.g., *relationship*, *existence*, *quarrelsomeness.*

Criterion #3: Quality and Number of Modifiers.   This criterion examines the semantics, morphology, and number of all modifiers within a single NP, but disregarding their syntax, counted separately under criteria #4 and #5. Criterion #3 considers the semantics and number of NP modifiers not, as noted, the syntactic structure of the NP construction. Modifiers included all types of determiners (*a*, *the*, *these*, *that*, *two*, *few*, etc.); adjectives (basic such as *little* or advanced such as *similar*, *extensive*, *industrial*); and lexical nouns inside the NP (e.g., the modifiers *step-dad* and *situation* in the NP *my step-dad's <u>ignorance</u> about the situation*). Here, again, the decision as to how many and what types of modifiers were assigned a particular score from 1 to 4 was language-dependent, since in this respect, too, English and Hebrew differ somewhat, including in the structure and distribution of nominal compounds (Berman, 2009; Berman & Clark, 1989). For example, a score of 1 would apply to both the English NP *a misunderstanding of me* since it has an article and a pronoun as modifiers, and also to its Hebrew counterpart *iy-havana klapay*, which has only a (bound) pronominal modifier in the prepositional phrase *klapay* 'toward-me.' A score of 2 was given to an NP like *very low tolerance for other people* since it has a basic adjective modified by an intensifier, and a noun modified by a determiner – and the same would be true for the Hebrew equivalent *sovlanut me'od nemuxa klapey anashim axerim* (literally 'tolerance very low toward people other'). A score of 3 applied to an NP like *a lack of acceptance for differences between people* = Hebrew *xoser kabala la-hevdelim beyn anashim*, since it has four modifiers (an article, two abstract nouns, and one categorial noun). And the highest score of 4 was assigned to the Hebrew NP *ha-derex ha-tova beyoter, ha-kcara beyoter o zo še-me'oreret paxot agresiviyut* 'the-way the-good the-most, the-short the-most or that which-arouses less aggressiveness = the best, the shortest, or the least aggression-arousing path,' with three modifying phrases, each of which is itself multiply modified.

Criterion #4: Syntactic Depth.   This refers to the number of complex governed nodes of NPs, PPs, and Ss contained within the target NP constituting the internal dependency architecture of the construction. Rank 1 is assigned to NPs with no internal depth, that is, that have no additional governed nodes beyond the initial head noun (e.g., <u>conflict</u>, *one <u>belief</u>*, *my other <u>friends</u>*). Rank 2 is given to NPs with two internal nodes, e.g., [*a natural <u>part</u>* [*of social interaction*]]; [*an internal <u>conflict</u>* [*between two*

opposing emotions]]; [*conflict* [*as a pure Platonic concept*]]; [*every* <u>realm</u> [*from the most internal and personal to the most public and universal domain*]]. Rank 3 is assigned to NPs with three internal nodes, e.g., [*one* <u>belief</u> [*that opposes* [*the belief* [*of another nation*]]]*; [*a* <u>lack</u> [*of harmony*] *and* [*agreement*] *between* [*forces*]]; while Rank 4 was assigned to the 'deepest' NPs, with four or more internal nodes, e.g., <u>people</u> [*that have conflicts* [*that disturb their peace of mind from the time* [*they are born until the day* [*they die*]]]] – which in fact turned out to be extremely rare across the sample. Note that in principle the criterion of syntactic depth applies similarly to both English and Hebrew, notwithstanding differences in the internal structure and ordering of the head nouns in the two languages.

Criterion #5: Syntactic Variability.   This final criterion refers to the number of *different* types of NP-modifying categories. Thus, a higher score is assigned to an NP with different categories of modifiers, e.g., relative clause, NP, PP, adjectives, infinitives, etc., on the principle that the more variety, the higher the score. The idea embodied in this criterion is that NP complexity derives not only from subordination but also from coordination and from the very variety of syntactic constructions within the NP. For example, the NP *a* <u>conflict</u> *with me* has only one type of modifier, the prepositional phrase *with me*, and so rates the lowest score of 1. The NP *one* <u>belief</u> *that opposes the belief of another nation* was assigned a score of 2 because it contains two different types of modifiers: the quantifier *one* and the relative clause *that opposes the belief of another nation* (which in itself contains two NPs *the belief* and *another nation* in a possessive construction). Also assigned a score of 2 on this criterion was the NP *very low* <u>tolerance</u> *for other people* since it has an adjective *low* and a prepositional phrase *for other people*. The NP *every* <u>realm</u> *from the most internal and personal to the most public and universal* rated a score of 3 since it contains the correlative coordinating construction of prepositional phrases *to . . . from* with two coordinated adjectives *most internal and personal* and *public and universal* in each. The (also very rare) score of 4 was given to variegated NP modification as in *many different kinds of extremely valuable contributions to the wellbeing of people living and working in Tel Aviv as the city of their choice.*

Results

The first analysis relates to syntactic complexity as measured by *mean clause length* (MCL). The clause, defined as 'a unified predicate describing a single situation (activity, state, or event)' (Berman & Slobin, 1994, pp. 660–663), is a critical unit for assessing text-embedded syntactic complexity, one that is directly relevant to NP analysis and so constituted the basic unit of analysis or 'textline' in our study. MCL was measured by mean number of words per clause, taking into account the different nature of what constitutes 'a word' in Hebrew versus English. For Hebrew, the seven prefixally attached closed class items that count as separate words in English – conjunctions, prepositions, and the definite article (Ravid, 2005) – were counted as part of the word they are attached to. Table 1 shows the mean number of words per clause by age, text type, and language.

    A four-way ANOVA of Age (4 groups: 9–10, 12–13, 16–17, and adults) × Language (2 languages: English, Hebrew) × Genre (2 genres: narratives, expositories) × Modality

Table 1. Mean Clause Length across Age Groups, Text Types, and Languages

Text Type	Age Group							
	9–10		12–13		16–17		Adults	
	English	Hebrew	English	Hebrew	English	Hebrew	English	Hebrew
Narrative spoken	5.95 (0.12)	5.84 (0.22)	6.82 (0.18)	5.6 (0.25)	6.02 (0.17)	5.72 (0.24)	6.24 (0.24)	5.71 (0.26)
Narrative written	5.83 (0.22)	5.5 (0.26)	5.82 (0.25)	5.57 (0.27)	6.18 (0.11)	6.92 (0.28)	5.6 (0.22)	6.24 (0.27)
Expository spoken	5.43 (0.34)	6.02 (0.34)	6.81 (0.19)	5.92 (0.29)	6.25 (0.28)	5.9 (0.21)	5.71 (0.27)	6.35 (0.32)
Expository written	5.06 (0.41)	5.81 (0.23)	5.94 (0.21)	6.11 (0.47)	6.84 (0.31)	7.68 (0.33)	6.7 (0.14)	8 (0.3)

Standard deviations in parentheses.

(2 modalities: speech, writing) was carried out on the data in Table 1. Results showed increase in MCL with age ($F(3,88) = 5.5$, $p < .003$), with a significant difference between the youngest group (aged 9–10) and the two older groups, and also that expository clauses are longer ($M = 6.28$) than narrative clauses ($M = 5.97$) (Genre ($F(1,88) = 4.12$, $p < .05$). A three-way interaction emerged for Age, Genre, and Modality ($F(3,88) = 2.83$, $p < .05$), with two main findings: first, differences in MCL between speech and writing were found only in expository texts, and only from high school up; and, second, age-related differences were found only in written texts. In addition, two 2-way interactions emerged: Age and Modality ($F(3,88) = 8.02$, $p < .001$); Genre and Modality ($F(1,88) = 7.09$, $p < .01$).[3] Beyond that, a Language and Modality interaction ($F(1,88) = 10.35$, $p < .003$) showed a discrepancy between spoken and written texts only in the Hebrew texts, while MCL was longer in written texts in Hebrew than in the written English texts.

Noun Phrase Length in Words

Table 2 shows the breakdown of scores on a scale from 1 to 4 for mean number of words per NP.

A four-way ANOVA (Age × Language × Genre × Modality) on the data in Table 2 showed increase in NP length with Age ($F(3,88) = 16.41$, $p < .001$), with adults differing significantly from the other three age groups. In addition, Hebrew was found to have longer NPs ($M = 1.45$) than English ($M = 1.38$) ($F(1,88) = 4.82$, $p < .04$), expository NPs to be longer ($M = 1.48$) than narrative NPs ($M = 1.36$) ($F(1,88) = 16.59$, $p < .001$), and written NPs to be longer ($M = 1.45$) than spoken NPs ($M = 1.39$) ($F(1,88) = 6.87$, $p < .02$). Two interactions mitigated these findings: Age, Language, and Genre ($F(3,88) = 3.26$, $p < .03$). One source of age-related differences derives from the expository texts in Hebrew, with a leap in mean NP length from grade school to junior high, and again between high school and adults; a second age-related difference derives from the

Table 2. Noun Phrase Length in Words across Age Groups, Text Types, and Languages

Text Type	Age Group							
	9–10		12–13		16–17		Adults	
	English	Hebrew	English	Hebrew	English	Hebrew	English	Hebrew
Narrative spoken	1.12 (0.12)	1.38 (0.22)	1.29 (0.18)	1.42 (0.25)	1.23 (0.17)	1.3 (0.24)	1.48 (0.24)	1.56 (0.26)
Narrative written	1.21 (0.22)	1.38 (0.26)	1.29 (0.25)	1.26 (0.27)	1.38 (0.11)	1.4 (0.28)	1.5 (0.22)	1.59 (0.27)
Expository spoken	1.39 (0.34)	1.37 (0.34)	1.23 (0.19)	1.47 (0.29)	1.41 (0.28)	1.38 (0.21)	1.64 (0.27)	1.54 (0.32)
Expository written	1.39 (0.41)	1.22 (0.23)	1.33 (0.21)	1.58 (0.47)	1.62 (0.31)	1.57 (0.33)	1.65 (0.14)	1.82 (0.3)

Standard deviations in parentheses.

division between adults compared with the three younger groups. The interaction of Age and Modality ($F(3,88) = 2.92$, $p < .04$) shows that written texts have longer NPs than spoken from high school onwards.

Semantic Abstractness of Noun Phrase Head

The third criterion of NP complexity analyzed the head noun of each NP on a scale of semantic abstractness from 1 to 4, as shown in Table 3.

A four-way ANOVA (Age × Language × Genre × Modality) on the data in Table 3 showed an increase in abstractness and categoriality of NP head with age ($F(3,88) = 46.96$, $p < .001$) with each age group differing from the one that precedes and follows it. In addition, Hebrew NP heads score higher ($M = 2.58$) than English ($M = 2.46$) ($F(1,88) = 7.85$, $p < .007$), expository NPs are more abstract ($M = 2.69$) than narrative NPs ($M = 2.35$) ($F(1,88) = 106$, $p < .001$), and written NPs are more abstract ($M = 2.57$) than spoken NPs ($M = 2.47$) ($F(1,88) = 29.15$, $p < .001$). Two interactions mitigated these findings. First, Language and Genre ($F(1,88) = 6.64$, $p < .02$), showing that in both languages expository texts reflect greater semantic abstractness than narratives, but Hebrew narratives contain more semantically abstract nominals than English narratives. Second, with respect to the interaction of Genre and Modality ($F(1,88) = 5.52$, $p < .03$), only expository texts show a difference in semantic abstractness between speech and writing, while both spoken and written expository texts show greater semantic abstractness than their narrative counterparts.

Quality and Number of Modifiers

Table 4 gives results for the criterion of Modifier Quality, which scored the amount and level of modifiers from a score of 1 for simplex NPs with merely an article or quantifier

Table 3. Semantic Abstractness of NP Head across Age Groups, Text Types, and Languages

Text Type	Age Group							
	9–10		12–13		16–17		Adults	
	English	Hebrew	English	Hebrew	English	Hebrew	English	Hebrew
Narrative spoken	1.95 (0.19)	2.17 (0.22)	1.96 (0.38)	2.38 (0.21)	2.38 (0.21)	2.5 (0.25)	2.58 (0.33)	2.69 (0.3)
Narrative written	1.99 (0.33)	2.21 (0.39)	1.99 (0.39)	2.47 (0.37)	2.45 (0.28)	2.58 (0.24)	2.71 (0.39)	2.62 (0.44)
Expository spoken	2.19 (0.28)	2.32 (0.25)	2.46 (0.31)	2.58 (0.19)	2.73 (0.33)	2.72 (0.25)	2.94 (0.23)	2.93 (0.19)
Expository written	2.3 (0.37)	2.34 (0.34)	2.6 (0.19)	2.74 (0.3)	3.01 (0.24)	2.95 (0.27)	3.11 (0.25)	3.03 (0.2)

Standard deviations in parentheses.

Table 4. Quality and Number of Modifiers across Age Groups, Text Types, and Languages

Text Type	Age Group							
	9–10		12–13		16–17		Adults	
	English	Hebrew	English	Hebrew	English	Hebrew	English	Hebrew
Narrative spoken	1.21 (0.12)	1.38 (0.22)	1.38 (0.24)	1.47 (0.24)	1.38 (0.13)	1.41 (0.25)	1.55 (0.19)	1.61 (0.24)
Narrative written	1.23 (0.2)	1.42 (0.15)	1.4 (0.22)	1.36 (0.23)	1.5 (0.08)	1.58 (0.27)	1.63 (0.23)	1.91 (0.27)
Expository spoken	1.43 (0.24)	1.46 (0.31)	1.25 (0.23)	1.56 (0.21)	1.55 (0.26)	1.5 (0.17)	1.7 (0.21)	1.7 (0.35)
Expository written	1.47 (0.35)	1.3 (0.33)	1.53 (0.19)	1.68 (0.37)	1.78 (0.29)	1.76 (0.29)	1.84 (0.19)	2.03 (0.34)

Standard deviations in parentheses.

to a score of 4 for high-level modifiers (e.g., derived adjectives or complex relative clauses).

A four-way ANOVA (Age × Language × Genre × Modality) on the data in Table 4 showed an increase in quality and number of NP modifiers with age ($F(3,88) = 27.61$, $p < .001$), with the two younger age groups clustering together as compared with high school students, and these differing from adults in nominal modification. In addition, Hebrew modifiers score higher ($M = 1.57$) than English ($M = 1.49$) ($F(1,88) = 6.75$, $p < .02$), expository NPs score higher on modifiers ($M = 1.6$) than narrative NPs ($M = 1.46$) ($F(1,88) = 24.66$, $p < .001$), and written NPs score higher on modifiers ($M = 1.59$) than spoken NPs ($M = 1.47$) ($F(1,88) = 33$, $p < .001$). The three 3-way interactions that emerged are particularly revealing: Age, Language, and Genre ($F(3,88) = 3.16$, $p < .03$)

Table 5. Syntactic Depth across Age Groups, Text Types, and Languages

Text Type	Age Group							
	9–10		12–13		16–17		Adults	
	English	Hebrew	English	Hebrew	English	Hebrew	English	Hebrew
Narrative spoken	1.14 (0.11)	1.29 (0.21)	1.29 (0.18)	1.28 (0.19)	1.23 (0.15)	1.22 (0.2)	1.37 (0.15)	1.48 (0.28)
Narrative written	1.17 (0.16)	1.34 (0.29)	1.24 (0.14)	1.25 (0.25)	1.35 (0.09)	1.43 (0.23)	1.48 (0.2)	1.75 (0.35)
Expository spoken	1.36 (0.3)	1.33 (0.31)	1.23 (0.22)	1.39 (0.23)	1.43 (0.23)	1.35 (0.15)	1.58 (0.19)	1.49 (0.33)
Expository written	1.37 (0.32)	1.21 (0.26)	1.38 (0.23)	1.47 (0.33)	1.58 (0.25)	1.62 (0.32)	1.62 (0.14)	1.87 (0.29)

Standard deviations in parentheses.

show that age-related differences in both languages center on expository texts, while differences between English and Hebrew in use of noun qualifiers are evident in narratives, as follows: in English, the narratives of the youngest and oldest groups differ in this respect, whereas Hebrew narratives show a difference only between the high school and adult groups. The interaction between Age, Language, and Modality ($F(3,88) = 3.67$, $p < .02$) shows that the only place where there is a difference in nominal modification between spoken and written texts within a given age group is in the Hebrew adults, with between-age differences occurring in both modalities in both languages. Finally, Age, Genre, and Modality ($F(3,88) = 2.73$, $p < .05$) shows that differences between written and spoken texts within the same age group are evident from junior high school on in both languages, but only in expository texts, which have the highest scores on number and quality of modifiers; whereas between-age differences are found in both narrative and expository texts. A two-way interaction of Age and Modality also emerged ($F(3,87) = 6.88$, $p < .001$).

Syntactic Depth

This criterion refers to the number of nodes subordinated to the head NP, scored on a scale from 1 to 4, as shown in Table 5.

A four-way ANOVA (Age × Language × Genre × Modality) on the data in Table 5 showed increase in syntactic depth of NPs with age ($F(3,88) = 20.96$, $p < .001$), with the two younger age groups again clustering together as compared with high school students, and these differing from adults. In addition, Hebrew NPs show more syntactic depth ($M = 1.42$) than English ($M = 1.36$) ($F(1,88) = 6.75$, $p = .05$), expository NPs score higher on syntactic depth ($M = 1.45$) than narrative NPs ($M = 1.33$) ($F(1,88) = 20.29$, $p < .001$), and written NPs score higher on syntactic depth ($M = 1.45$) than spoken NPs ($M = 1.34$) ($F(1,88) = 28.83$, $p < .001$). Two interactions mitigated these findings: Age, Language, and Modality ($F(3,88) = 3.06$, $p < .04$), as well as Age and Modality ($F(3,88) = 7.2$,

Table 6. Syntactic Variability across Age Groups, Text Types, and Languages

Text Type	Age Group							
	9–10		12–13		16–17		Adults	
	English	Hebrew	English	Hebrew	English	Hebrew	English	Hebrew
Narrative spoken	1.07 (0.09)	1.17 (0.16)	1.22 (0.15)	1.23 (0.22)	1.1 (0.1)	1.13 (0.21)	1.24 (0.16)	1.29 (0.15)
Narrative written	1.07 (0.13)	1.2 (0.24)	1.17 (0.18)	1.17 (0.16)	1.19 (0.06)	1.23 (0.28)	1.36 (0.26)	1.43 (0.32)
Expository spoken	1.16 (0.2)	1.2 (0.27)	1.11 (0.11)	1.24 (0.18)	1.22 (0.18)	1.18 (0.16)	1.36 (0.2)	1.29 (0.25)
Expository written	1.11 (0.18)	1.16 (0.24)	1.13 (0.1)	1.34 (0.33)	1.4 (0.22)	1.43 (0.28)	1.43 (0.15)	1.47 (0.32)

Standard deviations in parentheses.

$p < .001$). Only Hebrew adult texts differed in syntactic depth between speech and writing, and only in Hebrew were there differences between both the spoken and the written texts of high schoolers and adults. English showed more 'long-distance' age-related differences between the youngest group compared with the high schoolers and adults in syntactic depth of their NPs.

Syntactic Variability

The last criterion applied to assess the internal complexity of NPs analyzed the number of different types of NP modifiers, scored from 1 for only one type of modifier to 4 for highly varied modifiers, as shown in Table 6.

A four-way ANOVA (Age × Language × Genre × Modality) on the data in Table 6 showed an increase in syntactic variability of NPs with age ($F(3,88) = 14.31, p < .001$) with adults differing from all other age groups. In addition, Hebrew NPs show more syntactic variability ($M = 1.27$) than English ($M = 1.21$) ($F(1,88) = 5.06, p < .03$), expository NPs score higher on syntactic variability ($M = 1.27$) than narrative NPs ($M = 1.2$) ($F(1,88) = 8.8, p < .005$), and written NPs score higher on syntactic variability ($M = 1.27$) than spoken NPs ($M = 1.2$) ($F(1,88) = 19.09, p < .001$). One interaction of Age and Modality mitigated these findings ($F(3,88) = 7.43, p < .001$), showing that age-related differences between spoken and written texts were found only in the adult group, with the written texts of adults differing from all the schoolgoers, and with spoken texts showing a difference in syntactic variability of NPs between high schoolers and adults.

Further Analyses: Clausal Nominality

The results derived from applying the five criteria of NP complexity indicate that, as predicted, Hebrew is in some sense 'more nominal' than English. To substantiate this impression, we conducted a further analysis, independent of NP-internal complexity, to

Table 7. Clause Nominality: Total Number of Clauses Containing Two or More NPs, by Text Type and Language

# NPs per CL	Narrative Spoken		Narrative Written		Expository Spoken		Expository Written	
	Hebrew	English	Hebrew	English	Hebrew	English	Hebrew	English
Two	126	134	114	178	171	187	125	247
Three	21	13	32	20	55	32	48	61
Four	8	2	14	6	19	4	18	7
Five	5	1	2	–	6	–	14	3
Six	–	–	–	–	2	1	4	–
Seven	–	–	–	–	1	–	4	–

assess the relative contribution of NPs to overall text volume. Taking the clause as the next higher level of constituent structure, we examined the number of NPs per clause, regardless of their syntactic function (as subject, direct object, adverbial, etc.). Table 7 gives the breakdown in raw scores of number of clauses containing two or more NPs by text type and language across the participants.

The frequencies listed in Table 7 provide further support for our previous results regarding both text type and language. In text type, expository texts in general, and written expositories in particular, have more clauses with two or more NPs and, moreover, more of their clauses contain multiple NPs. In language, across all text types, Hebrew has more instances of clauses with three or more NPs per clause, and cases of clauses with five or more NPs are confined almost exclusively to Hebrew. This higher reliance of Hebrew on clauses with multiple NPs is even more apparent in the expository genre, applying even to the spoken expository texts in Hebrew. Note that this finding cannot be said to derive from differences in MCL, in which Hebrew and English are much the same (see Table 1).

Discussion

The study analyzed the structure and content of NPs as a relatively uncharted facet of developing text-embedded syntactic complexity. Three key themes emerged. First, there was an across-the-board age-related increase in NP complexity along with a rise in clausal complexity, with both developments being most marked in late adolescence. Second, this trend interacts with the discourse variables of genre and modality, with written expository texts emerging as more syntactically complex than oral texts, on the one hand, and than narratives, on the other – a pattern that was likewise most evident from high school up. Third, Hebrew showed greater NP complexity than English, with the patterns that emerged for age and text type being consistently more marked in Hebrew.

Educators and others familiar with children's language often comment that discourse produced by grade-schoolers manifests fewer words, shorter sentences, and less complicated structures than those of high school students. Overall, they have the impression

that children's language is not so rich as that of adolescents or adults. Our study shows that this intuition is well founded, through detailed analysis of one locus of this developmental progression. In general, NP complexity turns out to be related to amount and variety of noun modification. This facet of language structure and use places a heavy burden on younger children both because of memory constraints and because of their more limited knowledge base, which means they are less likely to use attributes in labeling entities in discourse. Optional modifying elements not only add to NP length, they also contribute to making texts more informative (Ravid & Berman, 2006). Reliance on elaborated lexical NPs emerges as a relevant means for evaluating the increased complexity of language use during the school years, in marked contrast to the largely pronominal and simplex nominal constructions occurring in early child language (Ravid et al., 2002). From middle childhood to adolescence and on to adulthood, NPs grow longer in words, they include more, and more varied types of modifiers, they reveal greater syntactic depth, and they employ semantically more abstract nouns as heads. Compare, for example, the following excerpts from the oral narratives in the Hebrew sample, with NPs underlined, elements that do not occur in Hebrew parenthesized in the glosses, and orthographically attached prefixes indicated by hyphens: *kol ha-xaverim sheli* *sixaku* 'all my friends played' (4th grade boy); *hi adayin xavera shelanu nora nora nora tova* 'she (is) still (an) awfully awfully awfully good friend of-ours' (7th grade girl); *tixnanu bemeshex shavua milxama beyn ha-cafon la-darom* 'we-planned in-the-course-of (a) week (a) war between the-north to-the-south' (11th grade girl); *pa'ar yaxasey ha-koxot ve-midat ha-hashka'a hem loh shavim* 'the-disparity-in balance-of power and-the-extent-of-investment are not equal' (female graduate student).

These quite typical examples are taken from oral narratives, which all our criteria reveal as lower on all scales of NP complexity than texts produced in writing, on the one hand, and in the expository genre, on the other. In contrast to the more abstract and generalized nature of reference in expository discourse, NPs in narratives often designate proper nouns – names of people and specific places or institutions – and so are less likely to elicit modification. In contrast, the finding that written expository texts constitute a favored site for use of complex NPs can be explained by the more monitored and stable, non-interactive nature of writing combined with the heavy cognitive demands of expository discourse for logical organization and elaboration of ideas. Grade-school children are as yet unable to avail themselves of what Olson (2006) terms the 'affordances' of formal written languages that characterize the 'literate language user' (Berman & Ravid, 2009). Nor do they have the broad knowledge base and extended experience with literacy-related activities required for using complex syntax in discussing abstract issues. These emerge in tandem with the dramatic cognitive developments that take place in adolescence, as a major period of transition toward adult-like autonomy, characterized by greater cognitive flexibility and more efficient processing (Paus, 2005), abilities that Steinberg (2005) attributes to the assembly of an advanced 'executive suite' of abilities.

Our results are also consistent with analyses of other linguistic systems in the same and related databases. Findings from numerous independent studies underscore the critical status of adolescence in the development of complex language use, and the important role of expository writing as a platform for the expression of a rich range of linguistic abilities in different languages. These include general measures of language productivity

(Ravid, 2004b; Ravid & Zilberbuch, 2003a, 2003b); nominal semantics (Ravid, 2006; Ravid & Avidor, 1998; Ravid & Cahana-Amitay, 2005; Ravid & Schiff, 2006); lexical usage (Nir-Sagiv et al., 2008); passive, middle-voice, and impersonal constructions (Berman & Nir-Sagiv, 2004; Jisa, Reilly, Verhoeven, Baruch, & Rosado, 2002; Jisa & Viguié, 2005; Ragnarsdóttir & Strömqvist, 2005; Tolchinsky & Rosado, 2005); linguistic register (Bar-Ilan & Berman, 2007; Ravid & Berman, 2009); and clause-combining syntax (Berman & Nir-Sagiv, 2007, 2008). In all these studies, the essay type of text, a written exposition discussing a socially relevant issue, elicited the most complex forms of linguistic expression in diverse domains. The fact that a significant cut-off point emerged between grade-school and junior-high school children, on the one hand, and high school students, on the other, underscores the special status of adolescence as a watershed in developing cognitive and communicative abilities.

Recall that, as detailed in the Methods section, we adopted a four-point scale in order to accommodate the full range of constructions elicited by our database for each of the five criteria. This analysis avoided the reductionism inherent in binary assessment, and made it possible to account for interacting factors within each criterion, such as both quantity and diversity of modifiers. We were further motivated by the need to take into account nominal constructions with long-distance dependencies of the type that have motivated much of contemporary linguistic theory over the past half-century. These are illustrated by the following examples from Chomsky and Miller (1958), with the heavy subject NPs underlined inside a sentence in each case: *The administrator who the intern who the nurse supervised had bothered* lost the medical reports; *The fact that the employee who the manager hired stole office supplies* worried the executive.

An unexpected and interesting finding was, however, that the four structural criteria – length in words, number and type of modifiers, syntactic depth, and syntactic variability – reflected a much narrower range than the possible low of 1 to a high of 4, scoring between 1 to 2 on average. One explanation is that this reflects the syntax of natural non-expert text production, as exhaustively analyzed here for NP constructions in 384 texts in two different languages. We might expect to find a broader range in expert writing and in texts that are more carefully edited than the unrevised, unplanned discourse that is considered here. In fact, however, a glance at expert expository texts in an English newspaper revealed considerable structural variety in the sense that just about every NP they contained was constructed differently from the others, while none of them in fact would achieve a top score of 3 or 4.

A more psycholinguistically motivated explanation lies not in the language users, but in the very nature of language use. We interpret our narrow-range finding as reflecting a cognitive constraint on the amount of processing that speaker-writers can apply in the course of non-monitored discourse construction. Besides, our analysis reveals syntactic complexity to derive from a cluster of factors, of which clause-linking subordination is only one. That is, as suggested by Chafe (1994), all the criteria involved in our analysis are subject to the limitations of efficient cognitive processing that constrain how long and complex an NP can be in actual use. For example, why did 2 turn out to be the upper limit on NP length? The processing window is known to be about five to seven words long. This corresponds to a score of 2 on both the Hebrew and English scales in our analysis, and means that NP size tends to conform to the size of the processing window

so as to enable parsing. Moreover, the number of hugely complex NPs with 10 words or so, which do occur sporadically in our sample, is limited – certainly within a given text and, it turns out, also across different language users. Relatedly, on the dimension of syntactic depth, there is a limit on how much novel information – encoded in PPs and relative clauses – can or should be reasonably packaged together into a single NP. Here, the limit is imposed not only by parsing constraints but also by how much information reader-hearers can tolerate in a single envelope.

Only the fifth semantic criterion of head noun abstractness averaged a higher score than 1 to 2, running from 2 to 3. This reflects command of an advanced lexical repertoire shown by older speaker-writers in our sample, and relates to the other, syntagmatic criteria of NP complexity in the sense that more abstract head nouns serve as a center of gravity attracting numerous and diverse modifiers (Ravid, 2006; Ravid & Avidor, 1998; Ravid & Cahana-Amitay, 2005; Ravid & Zilberbuch, 2003a, 2003b). However, semantically abstract nouns do not require the online parsing of structural components that is constrained by a narrow window of processing capacities. Besides, even quite young children recognize that they can raise linguistic register by using a word that is more lexically specific or less common (Ravid & Berman, 2009).

Cross-linguistically, NPs turned out to reveal largely similar developmental, cross-modal, and cross-genre trends in both languages. However, Hebrew showed greater NP complexity than English by four out of the five criteria we applied, and the patterns revealed by our analyses were consistently more marked in Hebrew. Moreover, a separate measure showed higher reliance on clauses with multiple NPs in Hebrew than English (Table 7), indicating that NPs take up relatively more 'clause space' in Hebrew. That is, although both languages reveal a similar number of words per clause across age groups and genres (Table 1), the way these words cluster together into phrasal constituents differs markedly, with a bias toward more nominal elements in Hebrew clause structure compared with English. Moreover, in writing, the locus of most complex NPs, Hebrew clauses are on average longer than those of English. The finding that clauses written in Hebrew contain more words than those written in English is remarkable, given that closed class items that constitute separate words in English (the conjunctions *and*, *that*, the definite article *the*, and four common prepositions) are – as noted in connection with Table 1 – written and hence were counted as part of the word to which they are attached in Hebrew (e.g., *xshe-xaver-ey-nu* 'and-as-that-friend-s-our = and when our friends').

It thus emerges that nominal elements play a more central role in the discursive syntax of Hebrew than English. Motivated by these convergent findings, we conducted an informal count of the number of words occurring in NPs in our database. We found that nearly half of all words across text types and age groups were taken up by words-in-NPs in Hebrew, as compared with only around one-third of the text volume on average in English, indicating that nominal elements occupy more 'text space' in Hebrew.

We attribute these differences to typological factors and the relative 'nominality' of Hebrew compared with English. This relates to a cluster of structural properties of Hebrew, including: (1) the occurrence of verbless clauses in a range of copular constructions (Berman, 1978) of the type particularly favored in expository texts (Berman & Nir-Sagiv, 2004); (2) the rich range of nominal constructions linking two nouns – including so-called 'construct state' bound compounds, periphrastic noun plus noun constructions linked by

the genitive particle *shel* or some other preposition, the complex highly literary 'double construct' linking, and denominal adjective modification (Berman, 1988, 2009; Ravid & Shlesinger, 1995; Ravid & Zilberbuch, 2003a, 2003b); and (3) the wide range of verb-related morphological patterns for deriving different types of nominalizations in Hebrew (Berman, 2004; Ravid & Avidor, 1998; Seroussi, 2004). Hebrew verb roots typically yield between three and eight different nominals, whereas English is far more restricted in this respect (cf. from the consonantal root *g-d-l* 'grow' – nouns like *godel* 'size,' *gidul* 'growth, crop,' *gdula* 'greatness,' *gdila* 'increase,' *gadlut* 'munificence,' *megadel* '(crop) raiser,' *migdal* 'tower,' *hagdala* 'enlargement,' *magdelet* 'magnifier'). Moreover, Hebrew is lacking in many features that contribute to the richness and length of the *verb phrase* in English: these include auxiliary and modal verbs (Berman, 1980a) as well as verb particles and other satellite elements which in Hebrew are typically encoded by adverbial prepositional phrases, as are manner adverbs, all of which involve nominal constructions.

The one criterion on which English and Hebrew NPs scored the same was syntactic depth, which included both phrasal and (relative) clausal subordination. We relate this to the fact that speaker-writers in both languages link clauses together to a similar extent, although they use different means to do so. Thus, Hebrew favors more linear, paratactic means of clause linkage as against the English preference for subordination, while on the other hand Hebrew relies on tensed relative clauses where English favors non-finite participial modification (Berman & Nir-Sagiv, 2008; Berman & Slobin, 1994).

In conclusion, NP complexity provides a valid diagnostic of developing syntax and of text construction abilities across genres and modalities. The same set of criteria can readily be extended to examine the variables of syntactic site (pre- or post-verbal) and function (subject, object of verb, object of preposition, etc.) in a text-embedded perspective. Production of complex NPs is a very late acquisition, which requires recruitment of a constellation of linguistic and cognitive processing resources and which reveals syntactic complexity as being of varied shape both within and beyond clause boundaries. By taking the clause as a basic frame of reference, we were able to show that clause complexity and density in words (mean length of clause) and in NPs (clausal nominality) increase in tandem with NP-internal complexity as measured by five independent criteria. Taken together, the patterns revealed by these different measures reflect the value of a multifaceted approach to a single construction type in the developmental psycholinguistics of discourse production.

Acknowledgments

Notes

1   The study thus disregards the discursive feature of overall 'N-slot lexicality,' computed as the proportion of lexical nouns as compared with pronominals or zero in a given text (Ravid et al., 2002).

2   The project, 'Developing Literacy in Different Languages and Contexts,' was funded by a Spencer Foundation major grant, Ruth Berman, PI, with Dorit Ravid responsible for Hebrew and Judy Reilly of San Diego State University for Californian English data collection.

3   Sources of interactions were investigated using the Bonferroni procedure. Differences reported were significant at the .05 level. Wherever a given three-way interaction contained a two-way interaction, sources of the interaction were examined only for the three-way interaction.

References

Andrews, D. A. (2007). The major functions of the noun phrase. In T. Shopen (Ed.), *Language typology and syntactic description*, *Vol. I, Clause structure* (2nd ed., pp. 132–223). Cambridge: Cambridge University Press.

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

Bates, E., & Goodman, J. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia and real-time processing. In G. Altmann (Ed.), *Language and Cognitive Processes*, *12* [special issue on the *Lexicon*], 507–586.

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

Berman, R. A. (1980a). On the category of auxiliary in Modern Hebrew. *Hebrew Annual Review*, *4*, 15–37.

Berman, R. A. (1980b). The case of an (S)VO language: Subjectless constructions in Modern Hebrew. *Language*, *56*, 759–776.

Berman, R. A. (1988). Language knowledge and language use: Binomial constructions in Modern Hebrew. *General Linguistics*, *28*, 261–285.

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

Berman, R. A. (1999). Relevant features of spoken and written Hebrew. In R. Aisenman (Ed.), *Developing literacy across genres, modalities, and languages* (pp. 43–64). Tel Aviv University: International Literacy Project, Working Papers Vol. II.

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* (pp. 9–34)*.* Amsterdam: John Benjamins.

Berman, R. A. (2005). Introduction: Developing discourse stance in different text types and languages. *Journal of Pragmatics*, *37*(2) [special issue on *Developing discourse stance across adolescence*], 105–124.

Berman, R. A. (2008). The psycholinguistics of developing text construction. *Journal of Child Language*, *35*, 735–771.

Berman, R. A. (2009). Acquisition of compound constructions. In R. Lieber & P. Stekauer (Eds.), *Handbook of compounding* (pp. 298–322)*.* Oxford: Oxford University Press.

Berman, R. A., & Clark, E. V. (1989). Learning to use compounds for contrast. *First Language*, *9*, 247–270.

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

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

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

Berman, R. A., & Nir-Sagiv, B. (2008). Clause-packaging in narratives: A crosslinguistic developmental study. In J. Guo, E. Lieven, S. Ervin-Tripp, N. Budwig, S. Özçalişkan, & K. Nakamura (Eds.), *Crosslinguistic approaches to the psychology of language: Research in the tradition of Dan I. Slobin* (pp. 149–162). Mahwah, NJ: Lawrence Erlbaum.

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

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

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

Biber, D. (1998). *Variations across speech and writing*. Cambridge: Cambridge University Press.

Boscolo, P. (1990). The construction of expository text. *First Language*, *10*, 217–230.

Britton, B. K. (1994). Understanding expository text: Building mental structure to induce insights. *Handbook of psycholinguistics* (pp. 641–674). New York: Academic Press.

Brown, R. (1973). *A first language.* Cambridge, MA: Harvard University Press.

Bruner, J. (1986). *Actual minds, possible worlds*. Cambridge, MA: Harvard University Press.

Bybee J., & Noonan M. (Eds.). (2002). *Complex sentences in grammar and discourse: Essays in honor of Sandra A. Thompson.* Amsterdam: John Benjamins.

Cameron, C. A., Hunt, A. K., & Linton, M. J. (1988). Medium effects on children's story rewriting and story retelling. *First Language*, *8*, 3–18.

Chafe, W. (1994). *Discourse, consciousness, and time: The flow of language in speech and writing*. Chicago, IL: Chicago University Press.

Chomsky, C. (1969). *The acquisition of syntax in children from 5 to 10.* Cambridge, MA: MIT Press.

Chomsky, N., & Miller, G. A. (1958). Finite state languages. *Information Control*, *11*, 91–112.

Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.

Cristofaro, S. (2003). *Subordination*. Oxford: Oxford University Press.

Diessel, H. (2004). *The acquisition of complex sentences*. Cambridge: Cambridge University Press.

Friedmann, N., & Novogrodsky, R. (2004). The acquisition of relative clause comprehension in Hebrew: A study of SLI and normal development. *Journal of Child Language*, *31*, 661–681.

Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure.* Chicago, IL: Chicago University Press.

Grimshaw, A. D. (2003). Genres, registers, and contexts of discourse. In A. C. Graesser, M. A. Gernsbacher, & S. R. Goldman (Eds.), *Handbook of discourse processes* (pp. 25–82). Mahwah, NJ: Lawrence Erlbaum.

Halliday, M. A. K. (1989). *Spoken and written language*. Oxford: Oxford University Press.

Hickmann, M. (2003). *Children's discourse: Person, space, and time across languages*. Cambridge: Cambridge University Press.

Hunt, K. W. (1965). *Grammatical structures written at three grade levels (Research Rep. No. 3).* Urbana, IL: National Council of Teachers of English.

Jackendoff, R. (1977). X-bar-syntax: A study of phrase structure. *Linguistic inquiry monograph 2.* Cambridge, MA: MIT Press.

Jespersen, O. (1909). *A modern English grammar on historical principles*. London: Allen and Unwin.

Jisa, H. (2004). Developing alternatives for indicating discourse stance. In D. Ravid & H. Bat-Zeev Shyldkrot (Eds.), *Perspectives on language and language development: Essays in honor of Ruth A. Berman* (pp. 357–374). Dordrecht: Kluwer.

Jisa, H., Reilly, J. S., Verhoeven, L., Baruch, E., & Rosado, E. (2002). Passive voice constructions in written texts: A crosslinguistic developmental study. *Written Language and Literacy*, *5*, 163–182.

Jisa, H., & Viguié, A. (2005). Developmental perspectives on the role of French *on* in written and spoken expository texts. *Journal of Pragmatics*, *37*, 125–142.

Khourajaina, E., & Tolchinsky, L. (2004). Discourse constraints on the lexical realization of arguments in Spanish. In R. A. Berman (Ed.), *Language development across childhood and adolescence* (pp. 83–109). Amsterdam: John Benjamins.

Loban, W. (1976). *Language development: Kindergarten through grade twelve* (*Research Rep. No. 18*). Urbana, IL: National Council of Teachers of English.

Longacre, R. E. (1996). *The grammar of discourse* (2nd ed.). New York: Plenum.

Lust, B., Foley, C., & Dye, C. (2008). The first language acquisition of complex sentences. In E. Bavin (Ed.), *Handbook of child language* (pp. 463–505). Cambridge: Cambridge University Press.

Lyons, J. (1977). *Semantics* (Vol. II). Cambridge: Cambridge University Press.

Marchman, V., & Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language*, *21*, 339–366.

Marchman, V., & Thal, D. (2005). Words and grammar. In M. Tomasello & D. I. Slobin (Eds.), *Beyond nature–nurture: Essays in honor of Elizabeth Bates* (pp. 139–164). Mahwah, NJ: Lawrence Erlbaum.

Mosenthal, P. B. (1985). Defining the expository discourse continuum. *Poetics*, *14*, 387–414.

Nir-Sagiv, B. (2008). *Clause packages as constructions in developing narrative discourse*. Unpublished doctoral dissertation, Tel Aviv University.

Nir-Sagiv, B., Bar-Ilan, L., & Berman, R. A. (2008). Vocabulary development across adolescence: Text-based analyses. In A. Stavans & I. Kupferberg (Eds.), *Studies in language and language education: Essays in honor of Elite Olshtain* (pp. 47–74). Jerusalem: Magnes Press.

Olson, D. (2006). Oral discourse in a world of literacy. *Research in Teaching of English*, *41,* 136–142.

Paltridge, B. (1997). *Genre, frames and writing in research setting*. Amsterdam: John Benjamins.

Paus, T. (2005). Mapping brain maturation and cognitive development during adolescence. *Trends in Cognitive Science*, *9*, 60–68.

Radford, A. (1990). *Syntactic theory and the acquisition of English*. Oxford: Basil Blackwell.

Ragnarsdóttir, H., & Strömqvist, S. (2005). The development of generic *maður/man* for the construction of discourse stance in Icelandic and Swedish. *Journal of Pragmatics*, *37*, 143–156.

Ravid, D. (2004a). 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.

Ravid, D. (2004b). 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* (pp. 337–355). Dordrecht: Kluwer.

Ravid, D. (2005). Hebrew orthography and literacy. In R. M. Joshi & P. G. Aaron (Eds.), *Handbook of orthography and literacy* (pp. 339–363). Mahwah, NJ: Lawrence Erlbaum.

Ravid, D. (2006). Semantic development in textual contexts during the school years: Noun scale analyses. *Journal of Child Language*, *33*, 791–821.

Ravid, D., & Avidor, A. (1998). Acquisition of derived nominals in Hebrew: Developmental and linguistic principles. *Journal of Child Language*, *25*, 229–266.

Ravid, D., & Berman, R. A. (2006). Information density in the development of spoken and written narratives in English and Hebrew. *Discourse Processes*, *41*, 117–149.

Ravid, D., & Berman, R. A. (2009). Developing linguistic register in different text types. *Pragmatics & Cognition*, *17*(1), 108–145.

Ravid, D., & Cahana-Amitay, D. (2005). Verbal and nominal expression in narrating conflict situations in Hebrew. *Journal of Pragmatics*, *37*, 157–183.

Ravid, D., Olshtain, E., & Ze'elon, R. (2003). Gradeschoolers' linguistic and pragmatic speech adaptation to native and non-native interlocution. J*ournal of Pragmatics*, *35*, 71–99.

Ravid, D., & Saban, R. (2008). Syntactic and meta-syntactic skills in the school years: A developmental study in Hebrew. In A. Stavans & I. Kupferberg (Eds.), *Language education in Israel: Papers in honor of Elite Olshtain* (pp. 75–110)*.* Jerusalem: Magnes Press.

Ravid, D., & Schiff, R. (2006). Morphological abilities in Hebrew-speaking gradeschoolers from two socio-economic backgrounds: An analogy task. *First Language*, *26*, 381–402.

Ravid, D., & Shlesinger, Y. (1995). Factors in the selection of compound-type in spoken and written Hebrew. *Language Sciences*, *17*, 147–179.

Ravid, D., & Shlesinger, Y. (1999). Modern Hebrew adverbials: Between syntactic class and lexical category. In E. Contini-Morava & Y. Tobin (Eds.), *Grammatical categories* (pp. 333–351). Amsterdam: John Benjamins.

Ravid, D., & Tolchinsky, L. (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: Speech and writing. *Written Language and Literacy*, *5*, 69–94.

Ravid, D., & Zilberbuch, S. (2003a). Morpho-syntactic constructs in the development of spoken and written Hebrew text production. *Journal of Child Language*, *30*, 1–24.

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

Scott, C. M. (2004). Syntactic ability in children and adolescents with language and learning disabilities. In R. A. Berman (Ed.), *Language development across childhood and adolescence* (pp. 111–133). Amsterdam: John Benjamins.

Seroussi, B. (2004). Hebrew derived nouns in context: A developmental perspective. *Phoniatrica et Logopaedica*, *56*, 273–290.

Slobin, D. I. (Ed.). (1985). *The crosslinguistic study of language acquisition. Volume I: The data.* Hillsdale, NJ: Lawrence Erlbaum.

Slobin, D. I. (2004). The many ways to search for a frog: Linguistic typology and the expression of motion events. In S. Strömqvist & L. Verhoeven (Eds.), *Relating events in narrative: Typological and contextual perspectives* (pp. 219–257)*.* Mahwah, NJ: Lawrence Erlbaum.

Steen, G. (1999.) Genres of discourse and the definition of literature. *Discourse Processes*, *28*, 109–120.

Steinberg, L. (2005). Cognitive and affective development in adolescence. *Trends in Cognitive Sciences*, *9*, 69–74.

Strömqvist, S., Nordvist, A., & Wengelin, A. (2004). Writing the frog story: Developmental and cross-modal perspectives. In S. Strömqvist & L. Verhoeven (Eds.), *Relating events in narrative: Typological and contextual perspectives* (pp. 359–394)*. Mahwah, NJ: Lawrence Erlbaum.

Tolchinsky, L., Johansson, V., & Zamora, A. (2002). Text openings and closings: Textual autonomy and differentiation. *Written Language and Literacy*, *5*, 219–254.

Tolchinsky, L., & Rosado, E. (2005). The effect of literacy, text type, and modality on the use of grammatical means for agency alternation in Spanish. *Journal of Pragmatics*, *37*, 209–238.

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

Van Valin, R. D. (2006). On the relationship between syntactic theory and models of language processing. In A. Bornkessel & M. Schlesewsky (Eds.), *Semantic role universals and argument linking: Theoretical, typological and psycho-neurolinguistic perspectives* (pp. 263–302). Berlin: Mouton de Gruyter.