

## RESEARCH

# Usage patterns in the development of Hebrew grammatical subjects

Elitzur Dattner, Liron Kertes, Racehli Zwilling and Dorit Ravid

Tel Aviv University, Tel Aviv, IL

Corresponding author: Elitzur Dattner ([dattner.elitzur@gmail.com](mailto:dattner.elitzur@gmail.com))

---

The grammatical subject is a multi-faceted linguistic notion embedded in morphology, syntax and discourse-pragmatics. In Hebrew, grammatical subjects are associated with two distinct word orders, differing along grammatical, semantic, and pragmatic axes. The study examines the growth and proliferation of Hebrew grammatical subjects in the spontaneous speech of preschool children in a Usage-Based perspective, taking into account inflectional, syntactic and semantic properties of the clause side by side with discursive information. The corpus used for this study consisted of the recordings of 54 children in six age groups from two to eight years, engaged in triadic peer talk. Subject, predicate morphology and word order were coded, and utterances were coded according to their conversational roles. Using cluster analysis, each age group was found to have a characteristic usage pattern of subjects with associated syntactic, semantic and discursive properties, underscoring the acquisition and development of grammatical subjects in Hebrew. The usage patterns emerging from the current corpus are taken as a manifestation of the Discourse Profile Constructions notion: probabilistic form-function correlations consisting of multiple sources of formal and functional information, pairing a usage pattern of clauses with a unified construal and discourse function. Three Discourse Profile Constructions emerged from the data: joint action planning, conversational narrative, and conversational presentation. Each of these was associated with a different patterning of lexical or pronominal subjects coupled with predicates with specific temporal features and different word order. These findings suggest that gaining command of the subject category is linked to communicative functions in development.

---

**Keywords:** Grammatical Subjects; Hebrew; Language acquisition; Morphology; Discourse; Conversation

---

## 1 Introduction

Syntactic development has been the topic of developmental psycholinguistic investigation since its inception (Brown 1973). The scope of what is considered as syntactic development has been broad, focusing in young speakers on inflectional morphology, *wh*-questions, subject-aux inversion (e.g., Santelmann et al. 2002; Goldberg 2006), head-complement order, and verb-argument structure (e.g., Tomasello 2003; McClure et al. 2006; Ambridge & Blything 2016). Studies on older, preschool children's syntax are few (Frizelle et al. 2018), and have mostly targeted the acquisition of relative clauses (Friedmann & Novogrodsky 2004; Arnon 2010; Kirjavainen et al. 2017; Haendler & Adani 2018). Researchers of both language development and processing are in agreement that learning how grammatical subjects are expressed and used is critical for young children, as subjects are related to virtually any aspect of syntactic and pragmatic development. Most previous studies have examined the realization of grammatical subjects (overt vs. null), mainly in dyadic contexts, at the early emergence of syntax (Borer & Wexler 1987; Bloom 1990; Valian 1990; 1991; Grinstead 1998; 2000; Levy & Vainikka 2000; Hacoen & Schaeffer 2007). Fewer

studies have focused on the acquisition and development of the category of grammatical subjects in the developmental psycholinguistic literature (Matthei 1987; Ferdinand 1996; Serratrice 2005; Veneziano & Clark 2016). The current study adopts a Usage-Based perspective (Tomasello 2009), examining patterns of usage in the development of grammatical subjects in the peer-talk of Hebrew-speaking children aged 2;0–8;0 years, taking into account morphological, syntactic and discursive factors.

### 1.1 The notion of subject

The subject position, one of the oldest grammatical relation concepts, is a core grammatical function, “a significant ingredient in many of the grammatical phenomena of the language, so that it seems plausible to recognize it as a structural primitive” (Andrews 2007). Considerable evidence from many languages supports this claim, as demonstrated, for example, by the fact that subjects are more prone to what is called wh-movement cross-linguistically than other elements (Keenan & Comrie 1977). However, it is also a concept that has challenged linguistic theories of all flavors, especially those keen on identifying grammatical universals (Marantz 1982; Chomsky 2000), “as most languages have something which looks like a subject, but subjects differ across languages in their nature and properties” (Falk 2006). In fact, the notion of a grammatical subject is considered one of those grammatical phenomena that can be fully captured only when elements from different levels of grammar interface (Lambrecht 1994). The nature of subjecthood is not only controversial in terms of grammatical theory, but also in terms of definition and applicable criteria (Allerton 1980; Van Valin & LaPolla 1997). These criteria involve different linguistic domains, such as agency in semantics, topic-hood in pragmatics, and case in morphology (Lambrecht 1994). This is especially important in view of typological studies of ergative languages, where the subject of an intransitive clause such as *The door slammed* behaves like the object of a transitive clause (e.g., the door in *John slammed the door*). These studies underscore the need to expand the notion of subject beyond that typical of so-called nominative or subject-prominent Indo-European languages (Li & Thompson 1976; Schachter 1976; Dixon 1994). Against this background, we find it useful to adopt the widely-used, theory-neutral, typologically motivated terminology (Comrie 1978; Dixon 1979), which defines three syntactic relations in the clause:

- (1) a. S: Sole argument of an intransitive clause, e.g., *The door slammed*
- b. A: Agent-like argument of a transitive clause, e.g., *John slammed the door*, denoting the role which is the most relevant to the accomplishment of the activity
- c. O: Patient-like argument of a transitive clause, e.g., *John slammed the door*, denoting the participant not coded as A.

Thus, subjecthood in nominative-accusative alignments is a grammatical relation expressing S and A, but not O. Two major phenomena indicate the existence of grammatical subjects (Andrews 2007): (i) overt, consistent structural features such as word order, case marking, and assignment of agreement; and (ii) less consistent, behavioral phenomena that are more abstract and less equivocal, usually relative to other grammatical phenomena. These indications are especially relevant to Hebrew, where two word orders prevail, with differing degrees of subjecthood (Ravid 1995; Melnik 2006).

### 1.2 Grammatical subjects in Modern Hebrew

Despite the massive syntactic changes Hebrew has undergone over its long history (Givón 1977; Bendavid 1988; Borer 1995; Shlonsky 1997; Hatav 2000; Doron 2005), its Semitic origins are still apparent in its rich morphology, including subject-verb agreement (Berman

1978; Ravid 1995; Schwarzwald 2002). While the current study focuses on syntactic phenomena, verb agreement is relevant here as grammatical subjects can be either lexical or pronominal, including pronominal omission. The issue of subject realization and omission in Hebrew is vastly studied in the literature, mainly in the context of pro-drop discussions. For example, it was shown to be related to accessibility (Ariel 1998), licensing (Borer 1989), and information structure (Melnik 2006). However, the prototypical behavior of subjects in Modern Hebrew can be generalized as shown in Tables 1 and 2. Tables 1 and 2 show the Hebrew agreement system regarding finite (past, present and future tense) main verbs, in terms of type of agreement (Table 1: number, gender, and person), and the options for pronominal omissions (Table 2).

Past and future verbs agree with their subjects in person, number and gender, while present tense verbs are marked only for number and gender agreement (Table 1). Pronominal subject omission is generally possible in 1<sup>st</sup> and 2<sup>nd</sup> person pronouns in past and future tense, and is prototypically not allowed in all other positions (Table 2). Hebrew may allow third person pronominal omission in some restricted contexts, such as adjunct subordinate clauses, with a particular semantic types of verbs (Melnik 2007), or in some accessibility related scenarios (Ariel 1998). In the following sections, we exemplify the main patterns of subject realization in Modern Hebrew.

### 1.2.1 Default word orders and subjects

Modern Hebrew has two default word orders, which differ along several grammatical and lexical axes and are highly relevant to the notion of subject (Berman 1980; Ravid 1995; Goldenberg 1998; Kuzar 2012). Framed in typological and discursive parlance, these orientations are two *alignments* (Comrie 1978; Dixon 1979), reflecting Dryer's (1997) SV vs. VS contrastive orders, which he regards as a fundamental property of language. One orientation is the Subject-First construction with A/S marked grammatical subjects followed by predicates on which they confer agreement. Another is the Predicate-First construction (discussed in Melnik 2006), which is close to the semi-ergative, or active alignments (Bubenik 1989; Du Bois 2017), with clause-final, non-canonical, sometimes O-like marked subjects.

### 1.2.2 The Subject-First construction

Two sub-constructions are actually included under the Subject-First construction, the habitat of typical grammatical subjects in Hebrew – the transitive AVO and the intransitive SV constructions. The first one, AVO, refers to transitive clauses typically starting with an

**Table 1:** Subject-verb agreement in Hebrew verbs according to temporal category.

	Person	Number	Gender
Future tense	✓	✓	✓
Present tense	×	✓	✓
Past tense	✓	✓	✓

**Table 2:** Pronominal omission in Hebrew verbs according to person.

	1 <sup>st</sup>	2 <sup>nd</sup>	3 <sup>rd</sup>
Future tense	✓	✓	×
Present tense	×	×	×
Past tense	✓	✓	×

agentive subject, followed by a lexical verb, in turn complemented by an object phrase, as demonstrated below in (2). The most typical object complement in AVO clauses is the direct object; either the indefinite unmarked object (2-a) or marked by the accusative marker *et*, preceding definite direct objects (2-b). Objects in AVO clauses can also be oblique or indirect, i.e., preceded by non-accusative prepositions, as in (2-c)–(2-d):

- (2) AVO clauses
- a. *ibadeti shney masrekim b-a-derex.*  
lost.1SG(A) two combs(O) on-the-way.  
'I lost two combs on the way.'
  - b. *ha-xatulim kar'u et ha-sapa.*  
the-cats(A) tore.3PL ACC the-sofa(O).  
'The cats tore the sofa.'
  - c. *hu xashav al ha-shiur.*  
he(A) thought on the-lesson(O).  
'He thought about the lesson.'
  - d. *hi azra le-dani la-kum.*  
she(A) helped.3.SG.FM to-dani(O) to-get up.  
'She helped Danny get up.'

The second Subject-First construction is the intransitive SV structure. SV constructions too start with grammatical subjects, followed by an intransitive predicate on which they confer agreement (see (3) below). SV predicates may contain an intransitive lexical verb, as in (3-a)–(3-b); or else they are copular constructions with grammatical verbs that agree with the grammatical subject (3-c):

- (3) SV clauses
- a. *dani saxa.*  
Dani(S) swam.MS.  
'Danny swam.'
  - b. *hitalaf-nu.*  
fainted-1.PL(S).  
'We fainted.'
  - c. *Rachel hayta xaxama.*  
Rachel(S) was.FM smart.FM.  
'Rachel was smart.'
  - d. *ha-delet nitreka.*  
the-door(S) slammed.FM.  
'The door slammed.'

In sum, Hebrew Subject-First constructions (2)–(3) host pre-verbal pronominal and lexical grammatical subjects, agreement-conferring NPs, expressing the range of thematic roles associated with the notion of subject (Van Valin 2004). They also express a non-constrained range of predicative content (Ravid & Cahana-Amitay 2005), both verbal and copular.

### 1.2.3 The Predicate-First construction

In contrast to Subject-First constructions, Predicate-First constructions lack canonical grammatical subjects (Berman 1980; 1990), and offer little marking of agreement (Ravid 1995; Preminger 2009; Haspelmath & Sims 2010). The NP slot occupied by grammatical subjects in this construction can only contain lexical subjects. They are also restricted in the range

of predicative content to non-dynamic or presentative construals (Ziv 1988; Melnik 2006; Dattner 2008). Two constructions are actually included under this Predicate-First label. First, existence/possession and modal/experience clauses expressing non-dynamic, mainly present-tense oriented content. These are demonstrated in (4) below.

- (4) Existence/possession and modal/experience clauses
- a. yesh li et ha-xacer haxi niflaa b-a-shxuna.  
be to.me ACC the-yard.FM(S) most wonderful.FM in-the-neighborhood.  
'I have the most wonderful yard in the neighborhood.'
  - b. carix le-calcel b-a-paamon.  
necessary to-ring in-the-bell.  
'You have to ring the bell.'

The second Predicate-First construction is an early-emerging class of presentative constructions (Berman 1985; Dromi & Berman 1986) with lexical verbs and subjects (5).

- (5) afa lo pitom cipor gdola.  
flew.FM to.him suddenly bird.FM(S) big.FM.  
'A big bird suddenly flew away from him.'

The fact that grammatical subjects do not constitute a homogeneous, coherent category in Hebrew poses a challenge to young speakers who base their learning on usage patterns in conversation (Fischer 2015). They have to learn about grammatical subjects from occurrences of two different word orders with distinct morphological, syntactic, semantic and discursive functions, where subjects make more and less canonical appearances. Therefore, the current study takes a developmental perspective on the characterization of the notion of *subject* in spoken Hebrew interaction (Fox & Thompson 1990; Berman & Slobin 1994; Dasinger & Toupin 1994; Helasvuo & Kyröläinen 2016).

### 1.3 Grammatical subjects in spoken interaction

A context-bound study that takes dialogical conditions into account has been proven relevant in many cases, especially in accounting for the emergence of pragmatically dependent elements (such as the emergence of pronouns, conversational contexts containing turn-taking, interaction, and third party reference). These, and other interactive phenomena utilize triadic or polyadic skills not used in dyadic contexts (Forrester 1988). Deixis forms, for example, may be learned through overhearing, and so does understanding of naming practices (Brenner 1983; Oshima-Takane 1988).

In adult-child conversation, adult interlocutors use scaffolding procedures to clarify and interpret children's utterances, while making the meaning of their own language forms accessible to the child by basing them in shared context or linking them to previous discourse. Thus, adults provide children with models of how their intentional meanings should find conventional expression (Murase et al. 2005; Howard et al. 2008).

The context of the current analysis is the natural conversation of Hebrew-speaking preschool and school children with their age peers, rather than with their caregivers. Peer talk offers a unique window on development, as children in peer interactions do not receive elaborate adult feedback facilitating linguistic communication (Blum-kulka et al. 2010; Schuele 2010). Therefore, peer interactions provide more realistic distributions of language structures and reliable information on linguistic usage (Hoff 2010; Veneziano 2010; 2018). Peer talk moreover renders children active participants in communicative events, tasked with expressing intentions and meanings to their interlocutors while taking into account

their discursive roles (Forrester & Cherrington 2009). As conversational agents, children participate in the construction of communicatively valid behaviors and are placed upfront for observing and treating those of the more experienced partner (Snow & Ferguson 1977; Gallaway & Richards 1994). Furthermore, utterances in conversation might have different roles. That is, an utterance can initiate a conversation (or a section within one), or it can function as a response to a previous utterance by an interlocutor. Utterances can also be structured based on previous utterances, thus resonating them (Du Bois 2007). These qualities of utterances provide children with different opportunities to manifest their involvement in conversation, taking into account that each type of utterance has a different function. We assume children's emerging conversational skills go hand in hand with grammatical acquisition (Serratrice 2005), making peer talk a particularly felicitous framework for studying the development of grammatical subjects.

#### **1.4 Seeking grammatical subjects in development**

Despite its obvious importance in the acquisition of syntax, the developmental literature does not contain an abundance of studies looking into the emergence and consolidation of grammatical subjecthood. Grammatical subjects make their appearance in child language investigation under three main umbrellas. One perspective, prevalent in the generative framework, focuses on subject slots as empty categories (De Villiers 1995), especially in relation to the agent and patient roles. According to de Villiers, the inventory of possible lexical NPs in the language is paralleled by a set of empty categories, which need to be learned (or parametrized) by the child, given the different syntactic and lexical properties in different languages. The acquisition problem according to this view is for the child to determine this inventory and the syntactic conditions under which empty categories are allowed. The omission of pronominal subjects by young children, termed *null subjects* (e.g., *want water*) has thus been a central topic of generative acquisition studies, explained as either due to children having different grammatical abilities than adults (Borer & Wexler 1987; Clahsen 1990; Valian 1991), or assuming an extended period where English allows null subjects (Hyams & Wexler 1993). Bloom (1993) strongly supports the parametric view, whereby children start off representing overt subjects as obligatory (as in English), but as a result of simple positive input – subjectless sentences in the speech of adults – children who are exposed to a null subject language can switch their grammars to pro-drop (as in Italian or Spanish, see Shin 2016) or topic-drop (as in Chinese). Finally, in Hebrew-specific perspective, Armon-Lotem (2008) regards the acquisition of the Hebrew finite verb as related to the morphosyntactic features of the subject of the sentence. She examines the emergence of verb morphology together with overt subjects in the speech of several Hebrew-speaking children from about 1;6 to 3 years of age. This study indicates that Hebrew number, gender and person marking first appears towards the end of the second year of life, while bare subjects first occur much earlier; however, no distributional evidence is offered.

A second perspective on learning grammatical subjects regards it as part of the acquisition of the verb's argument structure (Tomasello 1992; Gleitman & Gillette 1995), as a window on children's growing knowledge of verb semantics and tendency towards agentive structures (Matthei 1987; Brooks & Tomasello 1999; Alishahi & Stevenson 2008). These studies mainly focused on the number and type of verb complements across early development, frequency and meaning of specific constructions, and implications for learning (Bidgood et al. 2014; Ambridge & Blything 2016), including impaired populations (Fletcher & Ingham 1995). For example, Goldberg (2006) considers the verb as having more predictive value than the grammatical subject in generalizing argument structure constructions for learners. However, no study to date has systematically investigated the

evidence for the consolidation of grammatical subject structures and functions from early to the later preschool years.

A third approach to learning grammatical subjects focuses on children's early ability to detect the pragmatic and discourse conditions for subject realization and omission, as shown by, e.g., Allen (2000) for Inuktitut, and by Guerriero et al. (2001) for Japanese. One insight is that in languages with fixed word order, such as English, children rely on word order and pragmatic knowledge in comprehending active transitive sentences, whereas in languages with more flexible word order, the primary source of information about the agent/subject role is verb agreement (Abbot-smith & Serratrice 2015; Serratrice & Allen 2015). Important to our current topic, Paradis & Navarro (2003) show that the sources of information available to young native learners regarding subject realization at the pragmatics/syntax interface include both distributions of overt and null subjects in the ambient language as well as the specific events that children experience in on-going interaction. Paradis and Navarro show that, in line with other studies of Spanish and Catalan acquisition (Grinstead 1998), monolingual children in their study grasped the appropriate frequency and the functional discourse-pragmatics determinants of subject realization in Spanish at two years of age. This conclusion is shared by Salazar Orvig et al.'s 2010 study of clitic pronouns in dyadic conversations, who suggest that children's uses of pronouns reflect early pragmatic skills, acquired by the third year of life. They also suggest that the grammatical values of morphological devices are not acquired prior to their discursive value, but rather these pragmatic functions are associated with the grammatical level from the onset. Finally, peer talk studies of children aged 12–34 months demonstrate how children build and co-construct grammar dialogically in terms of the social-interactive projects they are trying to accomplish, negotiating how they stand vis-a-vis one another (Goodwin & Kyratzis 2012; Köymen & Kyratzis 2014).

Within these three perspectives to the development of grammatical subjects, previous research mostly examined isolated parameters. Consequently, research questions regarded the study parameter as a dependent variable, asking what affects its realization. For instance, the realization of the subject (as lexical vs. pronominal vs. zero, Serratrice 2005), or the development of a specific argument structure construction, can be studied in this way. However, engaging in language is an inherently multifactorial, context sensitive, social activity. That is, parameters do not evolve in isolation. Thus, the current paper proposes an integrated approach to the development of grammatical subjects, seeking for usage patterns that capture the contextual conditions for each type of grammatical subject in different constructions. In that we follow Croft's (2001) radical approach to grammatical phenomena, and suggest that the grammatical subject is not a unified notion. Rather, it is built bottom-up on a construction-specific basis. By revealing the subject-related usage patterns associated with each age group, we will be able to portray a picture of the developmental path through which such exemplar-based grammatical category is created, and a non-unified, probabilistic notion of grammatical subject is formed. Since the notion of *subject* is a multi-faceted phenomenon with relevance to morphological, syntactic and discourse-pragmatic information, we analyze its occurrence in children's peer-talk conversations across preschool ages, using cluster analysis to reveal multifactorial usage patterns. We expect to find usage patterns which are sets of utterances, grouped together according to statistically based within-cluster similarity and between-cluster dissimilarity. We hypothesize that the distribution of usage patterns in the corpus will not be uniform across age groups, such that each usage pattern will characterize a particular age range in the data. That is, for every age range there should be a usage pattern that is used significantly more frequently than others. Thus we will be able to characterize each age range in terms of its usage patterns of grammatical subjects. In order to so, we utilize the notion of Discourse Profile Constructions.

### 1.5 Form, function, and Discourse Profile Constructions

Language use is inherently contextual, concrete, and discriminatory (Ramscar et al. 2010). Form and function, therefore, are intimately related, in the most obvious way, especially in the case of grammatical form (morphology and syntax).<sup>1</sup> A verb in the past tense form, for example, has the function of construing a state of affairs that took place in a time before the time of speech. Such construal is mostly connected with narrative discourse. On the other hand, a verb in the future tense form has the function of discussing the future; that is, it is used to predict, plan, and convey intentions. In the same way, a subject in the first person plural form has the function of talking about the interlocutors as a unified entity, and a subject in the second person singular has the function of profiling the interlocutor as a topic in the current discourse. Thus, form is more than isolated, context-free morphology or syntax; form is always motivated by function, to the extent that morphemes, words, and sentences have meanings. That is, there is a functional (i.e., meaningful) difference between a first person plural pronominal subject and a first person singular subject, and there is a functional difference between a verb in the past form and a verb in the present form.

Such an approach to language use and language function as highly concrete stems from the Usage-Based approach to language, in which language knowledge, as well as linguistic phenomena, are the product of praxis. The Usage-Based view can be traced back to Wittgenstein who put forward the proposition that, for many cases, “the meaning of a word is its use in language” (Wittgenstein 1953). Usage, however, is a very fluid concept. Taking the notion of usage seriously, then, we need to adopt a model that is able to quantify the role of usage in building language knowledge. The Discourse Profile Construction hypothesis is such a model (Dattner 2019). Discourse Profile Constructions (henceforth, DPCs) are a manifestation of this view in accounting for the usage from which the meanings, or the functions, of grammatical constructs emerge. DPCs are emergent form-function correlations that consist of multiple sources of concrete formal and functional information, conventionally pairing a usage pattern of clauses with a unified construal or discursive function. A usage pattern in this context is a frequent co-occurrence of concrete elements of form *and* function (e.g., a first person singular subject and a past tense verb), realized as a cluster of tokens with high within cluster similarity and high between cluster dissimilarity. Moreover, Dattner (2019) shows that DPCs are extensions of the Construction Grammar notion of Argument Structure Construction (Goldberg 1995; Perek 2015). From a usage-based, discursive point of view, Argument Structure Constructions can be broadened to include concrete forms and functions, emphasizing the existence of *basic discursive scenarios* that correspond with concrete exemplars of usage patterns.

Note that other models try to answer the question of *what is a possible sentence in a language*, or “to what extent is the human faculty of language an optimal solution to minimal design specifications, conditions that must be satisfied for language to be usable at all” (Chomsky 2001). The approach taken in the present paper assumes that for modeling the grammar of a language the question should be *why does the speaker choose the concrete form A over form B to convey a construal of the world X*. Thus, the DPC hypothesis states that DPCs are the basic *clausal* form-function correlations in the language, constituting the main usage-based source from which grammatical meaning emerges. Consequently, the DPC hypothesis does not account for what *can* be done with language; rather, it accounts for what *is* done with language.

<sup>1</sup> Function, or meaning, can be abstract or discriminative in nature (Baayen & Ramsar 2015); this debate, however interesting, is not relevant to the present claim, since the DPC hypothesis can accommodate both conceptualizations of linguistic representation.



## 2 Data and method

### 2.1 The corpus

The peer talk corpus (Zwiling 2009) consists of a total of eight hours of recordings of 54 children in six age groups: 2;0–2;6, 2;6–3;0, 3;0–4;0, 4;0–5;0, 7;0–8;0, engaged in triadic conversations, three triads per age group. All participants were native Hebrew speakers from medium-high SES, with no language disorders or other developmental problems. Triadic conversation was selected for this study as requiring complex turn-taking, the usage of pronominal reference, and the comprehension of the role of a third person in conversation (Salazar Orvig et al. 2010). Each of the 18 triadic sessions was conducted in one of the children's homes, audio-recorded and transcribed. Each triadic session was 30 minutes long, except for the youngest group triads (ages 2;0–2;6), which were 40 minutes long in order to allow for a larger production of language. One triad in each age group was given instructions to celebrate a birthday to a puppet, while the other two triads were not given any instructions. The researcher did not engage in children's spontaneous interaction.

The peer-talk corpus comprises 36,490 words, in 11,870 utterances. Among these, the current paper examines a sub-corpus composed of only those utterances containing all pronominal (both overt and zero) and lexical grammatical subjects. This is the Subject Corpus (SC), containing 5,694 utterances in 11,283 words. Table 3 provides the number of words and utterances in the two corpora by age group.

### 2.2 Method

Every utterance was coded according to the following criteria. (1) Role in conversation: As noted above, conversational data allows one to explore whether (and how) speakers use different patterns in conversation which correlate with different functions. Thus, the following roles of utterances were coded: Initiated, Repeating, Response, or Resonating utterances. This coding was conducted separately by the first and second authors, with a discussion resolving disagreement. (2) Predicate temporality (tense): past, present, and future. Imperative and infinitive verbs were excluded as irrelevant to the current analysis of subjects. (3) Subject features: (i) lexical or pronominal (both overt and zero), (ii) person and number, and (iii) order relative to predicate. Table 4 provides a summary of the coding parameters.

#### 2.2.1 Revealing usage patterns using cluster analysis

Taking a multivariate exploratory statistics perspective, we sought usage patterns in an unsupervised statistical learning manner, specifically using Hierarchical Clustering on Principal Components (henceforth, HCPC, with the FactoMineR R package; Lê et al. 2008;

**Table 3:** Corpus size: number of words and utterances in the peer-talk corpus and the subject sub-corpus (SC).

Age group	Words	Utterances	Words in SC	Utterances in SC
2;0–2;6	2794	1381	717	400
2;6–3;0	4302	1465	1392	685
3;0–4;0	5765	1896	1840	935
4;0–5;0	6700	2071	1924	959
5;0–6;0	7362	2342	2405	1219
7;0–8;0	9387	2715	3005	1496
Total	36310	11870	11283	5694

**Table 4:** Coding parameters.

Parameter	Categories
Subject type	2 levels: Lexical, Pronominal (both overt and zero)
Person	3 levels: 1 <sup>st</sup> , 2 <sup>nd</sup> , 3 <sup>rd</sup>
Number	2 levels: Singular, Plural
Tense	3 levels: Past, Present, Future
Word order	2 levels: Pre-verbal, Post-verbal
Utterance type	4 levels: Initiated, Repeating, Response, Resonating
Total: 6	16

R Core team 2018). HCPC is a cluster analysis conducted on the principal components of a Multiple Correspondence Analysis (MCA; Divjak & Fieller 2014). Clustering based on the principal components of the MCA has two advantages. First, the MCA transforms categorical variables into a set of continuous variables (i.e., the principal components). Thus, it functions as a pre-processing step allowing to compute clustering on categorical data. Second, the MCA functions as a method for reducing noise in the data. Without such noise, the clustering analysis is more stable as it is done on the signal itself. In the present case, utterances were clustered together, groupings which we will describe according to the variables each utterance was coded for. We finally used Correspondence Analysis (CA) to reveal associations between age groups and usage patterns, as age was not part of the HCPC calculation. CA is based on a contingency table representing relationships between two variables, calculating  $\chi^2$  distances (Yelland 2010; Husson et al. 2017). For a comprehensive illustration of revealing usage patterns using cluster analysis, see Dattner (2015; 2019).

### 3 Results and discussion

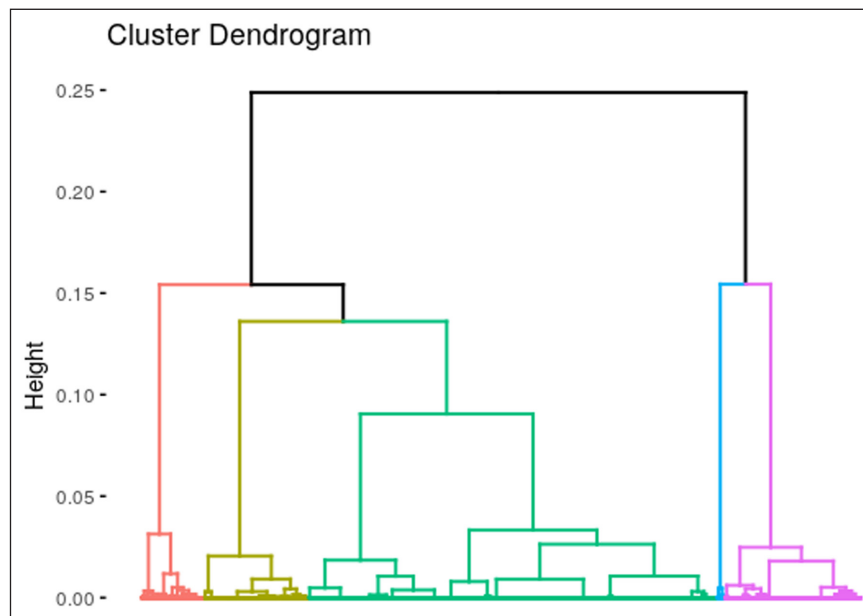
#### 3.1 Cluster analysis: Subject-related usage patterns

A Hierarchical Clustering on Principal Components analysis suggested an optimal division of the utterances in the data into five clusters. The optimal number of clusters was chosen according to growth of inertia, such that the difference in explained variation between dividing the data into four versus five clusters is greater than the difference between dividing the data into five versus six clusters. The cluster dendrogram is presented in Figure 1, and a description of each cluster according to the categories that were mostly associated with it is given below, summarized in Table 5.

Cluster one comprised 846 tokens, which were 14.85% of the 5,694 utterances in the data. It is a multifactorial usage pattern, mostly associated with initiated utterances containing a first person plural pronominal subject and a future tense verb (see Table 5 for a description of category distributions). It has the functions of (i) profiling the at-the-scene peer group as a single entity (hence the first person plural), and (ii) planning, suggesting and conveying intentions (hence the future tense verb). The following is a representative example for cluster one:

- (6) Cluster one:  
 bo nire et ze.  
 come will.see.1.PL ACC this.  
 ‘Let’s look at this one.’

Cluster two was the largest cluster of tokens in the data, comprising 3,054 utterances, constituting 53.63% of the utterances in the data. It is a usage pattern associated mostly



**Figure 1:** Hierarchical Clustering on Principal Components, colors indicating 5 clusters of the 5,694 sentences in the corpus: cluster 1 in olive-green, cluster 2 in green, cluster 3 in red, cluster 4 in light blue, and cluster 5 in purple.

with second person singular pronominal subjects, initiated and response utterances, with present tense verbs, and less so with first person singular. These characteristics are coupled with the functions of (i) highlighting the role of the interlocutor in the current state of affairs, and (ii) interpreting and commenting on current situations:

- (7) Cluster two:  
 ma at osa, Ofir?  
 what you.FM do, Ofir?  
 'What are you doing, Ofir?'

Cluster three consists of 475 tokens, which are 8.34% of the utterances in the data. This usage pattern is mostly associated with post-verbal lexical subjects, resonating utterances, and present tense verbs. Each of these forms is a marker of a particular function: The present tense verb marks the function of commenting on current state of affairs, and the post-verbal subject is used to introduce an out-of-the-scene entity to the discourse (hence the use of a lexical subject, see Ariel 2004).

- (8) Cluster three:  
 A: yesh lexa xitul.  
 be to.you diaper.  
 'You have a diaper.'  
 B: yesh lexa xitul gam.  
 be to.you diaper also.  
 'You also have a diaper.'

Cluster four was very small, consisting of 60 tokens, which are 1.05% of the utterances in the data. It was mostly associated with post-verbal lexical singular subjects in repeating utterances. This form, as was the case in cluster three, is linked with introducing a new entity into the discourse, as in the following:

**Table 5:** Cluster analysis: category distributions.

Cluster	Category	Rel freq corpus	Rel freq cluster	Global	v.test
one (14.85%)	tense=future	75.57	70.21	13.80	Inf
	number=pl	58.16	65.72	16.79	36.09
	person=1	32.52	84.75	38.73	29.98
	subject.type=pro	19.67	98.70	74.55	21.08
	word.order=pre	17.65	100.00	84.19	17.65
	utterance.type=initiated	17.14	78.13	67.74	7.21
two (53.63%)	word.order=pre	63.70	100.00	84.19	Inf
	subject.type=pro	71.94	100.00	74.55	Inf
	number=sg	63.02	97.77	83.21	33.58
	person=2	90.29	19.19	11.40	21.38
	tense=present	58.86	72.56	66.12	11.04
	tense=past	67.45	25.25	20.07	10.58
	utterance.type=response	64.95	27.18	22.44	9.27
	utterance.type=initiated	57.66	72.82	67.74	8.82
three (8.34%)	person=1	60.68	43.81	38.73	8.49
	utterance.type=resonance	95.19	100.00	8.76	Inf
	person=3	10.39	62.11	49.88	5.59
	tense=present	9.46	74.95	66.12	4.33
	word.order=post	11.89	22.53	15.81	4.01
four (1.05%)	subject.type=lex	10.56	32.21	25.45	3.45
	utterance.type=repeat	100.00	100.00	1.05	25.55
	word.order=post	2.22	33.33	15.81	3.36
	subject.type=lex	1.79	43.33	25.45	3.01
	person=3	1.41	66.67	49.88	2.61
five (22.11%)	number=sg	1.18	93.33	83.21	2.25
	word.order=post	85.89	61.40	15.81	Inf
	subject.type=lex	86.89	100.00	25.45	Inf
	person=3	44.33	100.00	49.88	Inf
	tense=present	26.69	79.83	66.12	12.02
	utterance.type=initiated	25.20	77.20	67.74	8.32
	number=pl	27.20	20.65	16.79	4.07

\* Each cluster is presented in Table 5 according to a descending order of the categories that characterize it the most. Thus, rows are not ordered in the same way across clusters. The table's columns are the following: The first column lists the clusters and their proportions in the corpus; the second column lists the specific categories linked to each cluster. The third column (Rel freq corpus) presents the relative frequency of tokens in the corpus that are coded for that category, which belong to the cluster. The fourth column (Rel freq cluster) indicates the relative frequency of the tokens in the cluster which possess each category. The fifth column (Global) lists the relative frequency of each category in the corpus, regardless of cluster. The last column (v.test) is related to the representation of the category in the cluster. If the v.test value is positive (or *Inf*, for infinite), the category is overrepresented in the cluster; if it is negative, the category is under-represented. Here we display only those categories with a positive v.test value, as these are the categories that are linked to each cluster.

- (9) Cluster four:  
 A: yesh po tiras.  
     be here corn cob.  
     ‘There’s a corn cob in here.’  
 B: yesh po tiras.  
     be here corn cob.  
     ‘There’s a corn cob in here.’

The usage pattern realized as cluster five consists of 1,259 tokens, constituting 22.11% of the utterances in the data. It is mostly associated with post-verbal (plural) lexical subjects, present tense verbs, in initiated utterances. The main difference between cluster five’s usage pattern and the usage pattern of clusters three and four is the role of the utterance in the conversation. While clusters three and four consists of resonating and repeating utterances (respectively), cluster five’s usage pattern is linked to initiated utterances. Such utterances constitute a more natural place in the discourse for introducing new entities, as in:

- (10) Cluster five:  
 beseder, yesh lanu po dapim.  
 fine, be to.us here pages.  
 ‘OK, here we have some sheets of paper.’

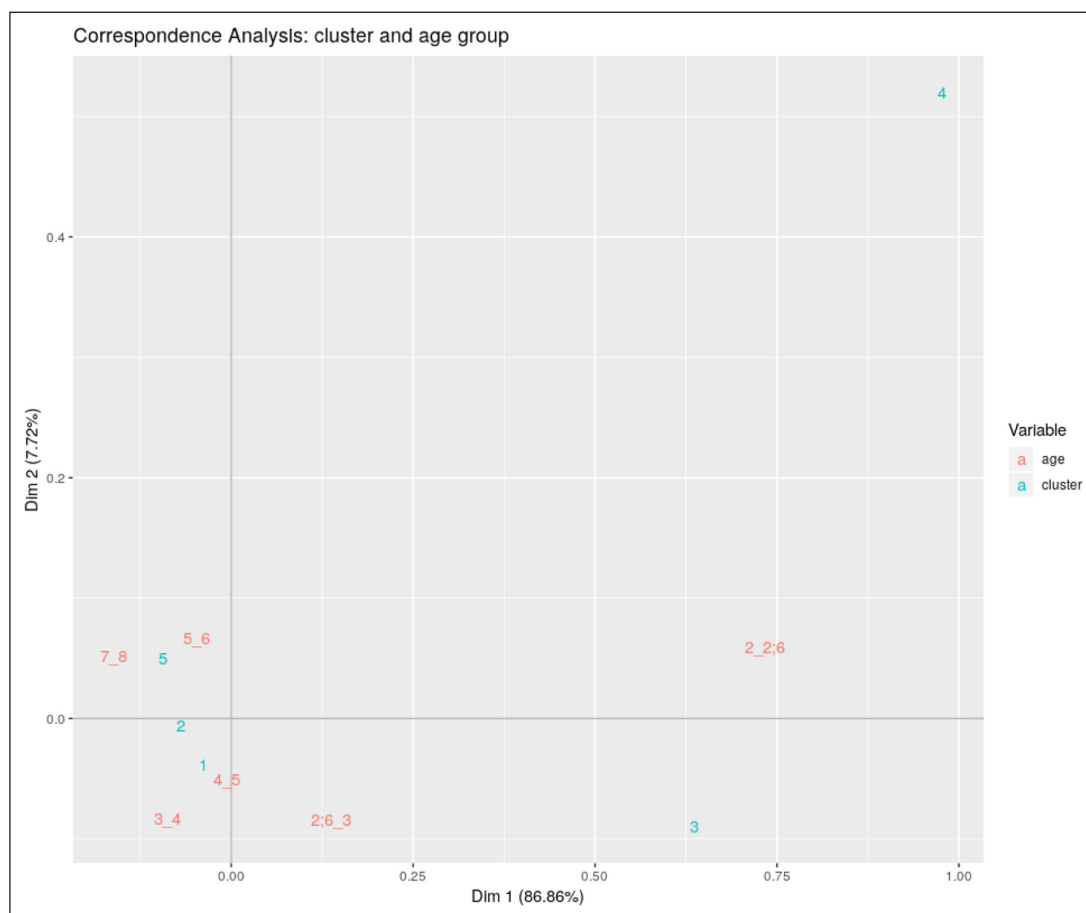
In the following section we will look at the distribution and correspondence of each usage pattern with regard to age groups.

### **3.2 Usage patterns in the development of the grammatical subject as a grammatical category**

The cluster analysis presented above was done without taking into account the association of each utterance with a specific age group. That is, for the purpose of the cluster analysis, the tokens in the data were considered to be a homogeneous corpus of 5,694 utterances, each coded for the same set of categorical variables. This allowed us to reveal usage patterns in the data regardless of the speaker’s age. However, given each token’s identification and cluster, we were able to conduct a post-processing analysis, looking for correspondences between cluster-based usage patterns and age groups. Figure 2 presents a Correspondence Analysis of age group and clusters.

The Correspondence Analysis map in Figure 2 depicts a clear developmental path in terms of similarity between age groups, expressed in shared cluster-based usage patterns of the grammatical subjects in the data. The youngest age group is located distantly relative to all other age groups, indicating that 2 year olds share very few usage patterns with the older children. Half a year older, the 2;6–3;0 children are more similar to the 3;0–4;0 and 4;0–5;0 age groups than to the 2–2;6 group. The 4;0–5;0 group is located a little higher on the map, due to its higher usage-based similarity to the older children in the corpus. And finally, the 5;0–6;0 and 7;0–8;0 groups are located near each other, with the 5;0–6;0 group functioning as a bridge between the 4;0–5;0 and the 7;0–8;0 age groups. That is, based on shared usage patterns of grammatical subjects in conversation, we can detect a gradient of three developmental stages in the preschool years: (i) 2;0–2;6 year olds, (ii) 2;6–5;0 year olds, and (iii) 5;0–8;0 year olds.

This developmental path can be further examined by the usage pattern that mostly corresponds to each group. Note, however, that “mostly corresponds” does not mean “exclusively used”: all usage patterns can be found in all age groups at some frequency. Rather,



**Figure 2:** Post-processing using Correspondence Analysis: clusters and age groups.

a characterization of an age group according to a usage pattern is based on a scalar, probabilistic distribution. The following sections present the developmental path of the grammatical subject usage patterns as realized by the correspondence analysis.

### 3.2.1 2;0–2;6 year olds

The youngest children in our database (2;0–2;6 years) are characterized by the usage patterns in clusters three and four (see Figure 2 and Table 5): post-verbal lexical subjects, in repeating and resonating utterances. The post-verbal position in Hebrew is mostly used to introduce and discuss new entities in the discourse (Melnik 2006). The fact that the 2;0–2;6 age group is linked to the usage patterns of clusters 3 and 4 indicates that they hardly engaged in a conversation, and produced few new grammatical structures without keeping a continued discourse topic. This is exemplified in the following excerpt (each turn represents a different speaker):

- (11) A: hine ktanim.  
 here small.PL.  
 ‘Here are some small ones.’
- B: hine ktanim.  
 here small.PL.  
 ‘Here are some small ones.’
- A: tiri.  
 look.FM.  
 ‘Look.’

- B: ze shaxor.  
this black.  
'This is black.'
- A: hine cahov.  
here yellow.  
'Here's(a) yellow one.'
- B: hine shaxor.  
here black.  
'Here's a black one.'
- A: hine cahov.  
here yellow.  
'Here's (a) yellow one.'

Moreover, looking at the conversations in this age group in a qualitatively perspective, we can see that the youngest children in our corpus tended to re-use given grammatical structures (realized as resonating and repeating utterances), in order to continuously introduce concrete, extant entities denoted by their size and color to the conversation in skeletal clauses devoid of verbs, as in the following examples:

- (12) Repeating utterance  
A: yesh po tiras.  
be here corn cob.  
'There's a corn cob in here.'
- B: yesh po tiras.  
be here corn cob.  
'There's a corn cob in here.'
- (13) Resonating utterance  
A: yesh lexa xitul.  
be to.you diaper.  
'You have a diaper.'
- B: yesh lexa xitul gam.  
be to.you diaper also.  
'You also have a diaper.'

### 3.2.2 2;6–5;0 year olds

The next age groups that share similar usage patterns according to the CA analysis can be characterized by the usage patterns of cluster three – resonating utterances in the present tense and post-verbal lexical subjects (also characterizing the younger children); and cluster one – initiated utterances containing a first person plural pronominal subject, and a future tense verb. (14) and (15) below exemplify the usage patterns of cluster three and one in these age groups, respectively:

- (14) A: hine uga.  
here cake.  
'Here's a cake.'
- B: hine hamon ugot.  
here a lot of cakes.  
'Here are a lot of cakes.'

- (15) bo nire et ze.  
 come will.see.1.PL ACC this.  
 'Let's look at this one.'

These examples indicate how the peer talk of children between the ages of two and a half to five years differs from that of their younger peers. First, they used nominal labels for new entities rather than their tangible characteristics (e.g., *cake* rather than *the small one*). Moreover, their clauses tended to include finite verbs and inflected nominals in addition to the skeletal verbless clauses used by younger participants. The 2;6–5;0 year olds already used resonating rather than repeating utterances, and a less restricted use of grammatical subjects, mainly including first person plural. That is, they re-use accessible structures to produce new and extended verbalizations, and they initiate conversations with reference to the group of interlocutors (1<sup>st</sup> person pl.), mainly discussing mutual planning and intentions in collaborative future tense, as in (15) above, and (16) below. These examples indicate the emergence of conversations.

- (16) A: bo nistakel kan b-a-tmunot.  
 come will.look.1.PL here in-the-pictures.  
 'Let's look at the pictures here.'
- B: ani marsha.  
 I allow.FM.  
 'I agree.'
- A: anaxnu lo yexolim la-kaxat.  
 we not able to-take.  
 'We can't take [it].'
- B: oy, ze kimat nofel.  
 oy, this almost falls.  
 'Oy, it's almost falling.'
- A: bo'u.  
 come.PL.  
 'Come.'
- B: ze omed li-pol.  
 it stands to-fall.  
 'It's going to fall.'
- A: ze omed li-pol.  
 it stands to-fall.  
 'It's going to fall.'

While we defined this age range uniformly (2;6–5;0), we can nevertheless detect a developmental path within this range, albeit a milder one. The 3;0–5;0 groups are less structurally restricted in terms of characterization by a specific cluster, as represented by the distance from the 2;6–3;0 group on the CA map in Figure 2. Unlike the younger peers, the 3;0–5;0 year olds also *initiate* conversations using new structures rather than producing resonating utterances. In (17), an interlocutor from the 3;0–4;0 group is addressed explicitly by name in a well-formed question, responding with an inflected verb describing the act of handing out tableware for a planned meal. In the exchange that follows, the speakers use the predicate-first construction to describe the current situation, but contrary to the youngest age group, this usage targets explicitly labeled objects (*glasses* and *plates*), includes a logical conjunction (*because*), and already demonstrates the characteristic absence of agreement between the predicative adjective (*missing*) and the post-verbal subject.



- (17) A: ma at osa, Ofir?  
 what you.FM do, Ofir?  
 ‘What are you doing, Ofir?’
- B: mexaleket.  
 handing out.FM.  
 ‘Handing out.’
- A: xaser kos.  
 missing.MS.SG glass.FM.SG.  
 ‘A glass is missing.’
- B: xaser calaxot.  
 missing.MS.SG plates.FM.PL.  
 ‘Plates are missing.’
- A: ki yesh od kos axat.  
 because be more glass one.  
 ‘Because there’s one more glass.’

In the CA map (Figure 2) we can see that the 3;0–4;0 and 4;0–5;0 groups are also characterized by the use of cluster two, which is less linked with the 2;6–3;0 group. Cluster two contains initiated and response utterances, with first and second person singular pronominal subjects in past tense. Example (18) illustrates this usage in a child’s commentary on her interlocutor’s hypothetical action, and example (19) is a response describing a hypothetical occurrence:

- (18) be-keilu at lakaxt et ze.  
 in-as if you.FM.SG took.2.FM.SG ACC this.  
 ‘Let’s pretend you took it.’
- (19) lo, nisharti po im ima.  
 no, stayed.1.SG here with mommy.  
 ‘No, I stayed here with mommy.’

### 3.2.3 5;0–8;0 year olds

The oldest age range is mostly characterized by cluster five’s usage pattern: initiated utterances in the present tense, with post verbal (plural) lexical subjects:

- (20) beseder, yesh lanu po dapim.  
 fine, be to.us here pages.  
 ‘OK, here we have some sheets of paper.’

The groups in this age range are located at the top half of the CA map due to their use of post verbal lexical subjects, shared with the youngest group. However, while the young children used this structure in repeating (cluster 4, see example (12) above) or resonating (cluster 3, see example (13)) utterances, the older groups are using it to appropriately introduce new entities (both concrete/present and abstract/absent) into the discourse in initiated utterances (Examples (20) and (21)):

- (21) tihye mesiba nehederet.  
 will.be.3.FM.SG party.FM great.FM.  
 ‘There will be a great party.’

While distinguished by cluster five, the proximity of the older age groups to the 3;0 to 5;0 age groups indicates that the former too use the common patterns of clusters 1 and 2:

initiated and response utterances with first person (plural and singular) and second person pronominal subjects, in the past, present and future tenses (as in (22), and (25)–(26)):

- (22) axshav nikax balonim.  
 now will-take.1.PL balloons.  
 ‘Now we’ll take balloons.’

That is, the oldest age range in the current data shows a non-restricted use of the grammatical subject category in conversation, appropriately utilizing specific usage patterns for the task they are most suitable for.

### 3.3 Usage patterns as Discourse Profile Constructions

Pairings of usage patterns and discourse functions, like the pairings revealed in our results, are described by Dattner (2015; 2019) as constituting Discourse Profile Constructions. Discourse Profile Constructions (DPCs) are emergent form-function correlations that consist of multiple sources of formal and functional information, conventionally pairing a usage pattern of clauses with a unified construal or discursive function. DPCs are claimed to be the basic clausal form-function correlations in the language, and to constitute the main usage-based source from which grammatical meaning emerges (Dattner 2019).

Three DPCs emerged from the data regarding the use of grammatical subject in conversation. Note that the labels for the DPCs are ad-hoc in that they are built bottom-up. That is, it is not the case that there is a predefined inventory of DPCs in the grammar of the language, such that the child needs to acquire it. Rather, DPCs are probabilistic clusters of similar usage events. Thus, only if a set of tokens are linked to a single discursive function frequently enough, and this function has discriminatory and discursive importance, a DPC will emerge. Importantly, these DPCs are relevant to a conversational type of linguistic interaction. In other types of linguistic data (e.g., in written expository texts, or in child directed speech by parents) different DPCs might emerge with regard to the use of grammatical subjects, that will join the heterogeneous category of subjects in Hebrew.

The first DPC is the *joint action planning DPC*, emerging from the tokens of cluster one: initiated utterances with a pre-verbal, first person plural pronominal subject (S/A) and a future tense verb, as in:

- (23) bo’u na’ase hacba’a.  
 come.2.PL will.do.1.PL voting.  
 ‘Let’s vote.’

This is a pattern of grammatical features (constituting both form and function), that serves a discursive function within conversation: it is used to plan joint actions, with reference to the interlocutors as a unified group of agents. According to the present corpus, this DPC is frequently used only from age 3;0 and up.

The second DPC is the basic *conversational narrative DPC*, emerging from the tokens in cluster two: pre-verbal 1<sup>st</sup> and 2<sup>nd</sup> person singular pronominal subjects (S/A), in present and past tense, in both initiated and response utterances:

- (24) A: Amit, at yexola le-haxzik et ha-bama kaxa?  
 Amit, you.FM can.FM to-hold ACC the-stage like this?  
 ‘Amit, can you hold the stage like this?’  
 B: ani lo yexola kol ha-hacaga.  
 I not can.FM all the-show.  
 ‘I can’t do it throughout the show.’

This pattern is used to serve the communicative function of describing current actions, commenting on past events, and telling narratives, while identifying each of the interlocutors as an agent and keeping continuous conversational topics. That is, it is a usage pattern composed of concrete usage events in which these grammatical features and functions corresponded with each other. The central location of this usage pattern on the CA map in Figure 2 is an indication to its wide use by all age groups in the data, except for the youngest 2 year olds. That is, it is a basic conversational pattern of initiation and response, concerning the interlocutors themselves (in 1<sup>st</sup> and 2<sup>nd</sup> person singular), for commenting and narrating in present tense, or a hypothetical past tense (showing a well developed inflectional system and the use of adverbials such as *suddenly* or *one moment* to enhance the pretend situations):

- (25) be-kei'lu at lakaxt et ze, ve-lo hersheti, ve-pit'om  
 in-as if you.FM.SG took.2.FM.SG ACC it, and-not allowed.1.SG, and-suddenly  
 amarti lax asur.  
 said.1.SG to-you forbidden.  
 'Let's pretend that you took it, and I didn't allow it, and suddenly I told you "no".'
- (26) A: bo'i.  
 come.2.FM.SG.  
 'Come.'
- B: rega, ani sama xulca.  
 moment, I put.on.1.FM.SG shirt.  
 'One moment, I'm putting on a shirt.'

The third DPC identified here is the *conversational presentation DPC*, emerging from the tokens of cluster five: post-verbal lexical subjects (mainly S) with verbs in the present tense:

- (27) li, le-Ana'el ve-le-Ofri yesh manuy le-hacagot.  
 to.me, to-Ana'el and-to-Ofri be subscription to-shows.  
 'I, Ana'el, and Ofri have a theater subscription.'

This pattern is used to introduce new topics to the discourse and to refer to less accessible entities outside of the conversational arena. Importantly, we show the learning curve of this DPC: The youngest children use it mostly with repeating utterances (cluster 4). That is, they are not introducing new entities, but rather *repeating* the presentation of such by their peers (as in (12) above). The mid-range children use this DPC mainly in *resonating* utterances (cluster 3). That is, while not repeating, they still re-use an immediately available structure:

- (28) A: yesh li gam et ha-kadur ha-ze.  
 be to.me too ACC the-ball the-this.  
 'I also have this ball.'
- B: yesh lexa gam et ha-sha'on.  
 be to.you too ACC the-watch.  
 'You also have the watch.'

Finally, the older groups appropriately use this DPC in initiated utterances (cluster 5). That is, once presented to the discourse, an off-stage entity (the star in (29) below) becomes accessible and is further referred to using a pronominal object.

- (29) A: sheli kvar ne'ebad ha-koxav.  
mine already was.lost.3.MS.SG the-star.  
'My star is already gone.'
- B: ulay kvar nimca oto.  
maybe already will.find.1.PL ACC.it.  
'Maybe we'll find it.'

Summing up the discussion of our results, we see that each usage pattern serves a particular communicative function. Importantly, note that each function can potentially be served by other grammatical features as well. However, we show that each function is prototypically served by a single usage pattern, which is a combination of morphological, syntactic, semantic and pragmatic information. This is a concrete, discourse-oriented, probabilistic view of the links between form and function in language, made possible by assuming a similarity based categorization of grammatical notions realized as Discourse Profile Constructions.

The associations between DPCs and age groups demonstrate the development of the grammatical subject category as a heterogeneous, exemplar-based category (Abbot-Smith & Tomasello 2006) vis-a-vis conversational skills: At the young age of 2;0–2;6 years, Hebrew-speaking children hardly relate to their interlocutors, and hardly keep track of a coherent topic, hence the use of the usage patterns of clusters 3 and 4. That is, they do not engage in conversation. Conversation begins to sprout at the age of 2;6 years according to the present corpus, with the mid-range ages referring mainly to themselves and their interlocutors as subjects (cluster 1, and to some extent cluster 2). Full command of conversation appears only in the older groups of 5;0–8;0 years old, realized as a non-restricted use of usage patterns, and command of all types of subjects within the grammatical subject category. The older children in the corpus talk about real as well as hypothetical events, in which both them or their interlocutors may be realized as subjects, as well as entities which are not part of the speech event itself. This suggests a less restricted discourse, containing reference to the interlocutors and to the off-stage scenario, presenting entities to the discourse and maintaining them, initiating conversations as well as responding, planning, commenting, and telling narratives. The present study thus shows that distributional information of multiple factors is important for categorization, facilitating the learning of the grammatical subject category (c.f. Romberg & Saffran 2010).

#### 4 Conclusion

Du Bois, Kumpf, and Ashby (2003) note that “[t]o understand grammar, find out how it is used.” In order to account for the development of the grammatical subject in Hebrew, the present study shows how it is used in conversation, in the peer talk of children 2–8 years old. We show that seeking a homogeneous abstract notion of *grammatical subject* is not a felicitous approach; rather, we identify three Discourse Profile Constructions of subjects in Hebrew, corresponding to three different usage patterns and three discursive functions. We show that these DPCs are not evenly distributed across age groups, but rather that each pattern is correlated with a particular age range, ordered according to a developmental path of conversational and grammatical skills.

Language can be metaphorically conceptualized as a toolbox consisting of unique formal devices used to convey a message, discriminating a particular construal of a state of affairs from all other possible construals. As tools are extensions of the body, enabling it to accomplish tasks it is not physically sufficient for, so do linguistic usage patterns to the mind. Language development, thus, is learning to properly use each linguistic device for the purpose it is most suitable. Continuing with the toolbox metaphor, a novice user

might use a hammer for driving both a nail and a bolt into a wall, while a professional user will use a hammer for the former and a screwdriver for the latter. That is, a temporary equilibrium state of learning is reached when a user (a speaker in our case) uses a variety of tools for a variety of tasks. The present paper shows that different patterns of linguistic forms conventionally correspond with different conversational functions to form Discourse Profile Constructions, and that young children use a small set of patterns for all tasks at hand, while older children seem to have mastered a wider range of tools, better fitting each usage pattern for a different task.

### Abbreviations

ACC = accusative, CA = Correspondence Analysis, DPC = Discourse Profile Construction, FM = Feminine, HCPC = Hierarchical Clustering on Principal Components, MCA = Multiple Correspondence Analysis, MS = Masculine, PL = plural, SG = singular

### Competing Interests

The authors have no competing interests to declare.

### References

- Abbot-Smith, Kirsten & Ludovica Serratrice. 2015. Word order, referential expression, and case cues to the acquisition of transitive sentences in Italian. *Journal of Child Language* 42(1). 1–31. DOI: <https://doi.org/10.1017/S0305000913000421>
- Abbot-Smith, Kirsten & Michael Tomasello. 2006. Exemplar-learning and schematization in a usage-based account of syntactic acquisition. *The Linguistic Review* 23(3). 275–290. DOI: <https://doi.org/10.1515/TLR.2006.011>
- Alishahi, Afra & Suzanne Stevenson. 2008. A computational model of early argument structure acquisition. *Cognitive Science* 32(5). 789–834. DOI: <https://doi.org/10.1080/03640210801929287>
- Allen, Shanley E. 2000. A discourse-pragmatic explanation for argument representation in child Inuktitut. *Linguistics* 38(3). 483–521. DOI: <https://doi.org/10.1515/ling.38.3.483>
- Allerton, David J. 1980. Grammatical subject as a psycho-linguistic category. *Transactions of the Philological Society* 78(1). 40–61. DOI: <https://doi.org/10.1111/j.1467-968X.1980.tb01032.x>
- Ambridge, Ben & Ryan P. Blything. 2016. A connectionist model of the retreat from verb argument structure overgeneralization. *Journal of Child Language* 43(6). 1245–1276. DOI: <https://doi.org/10.1017/S0305000915000586>
- Andrews, Avery D. 2007. The major functions of the noun phrase. In Timothy Shopen (ed.), *Language typology and syntactic description (second edition)* 1. 132–223. Cambridge, New York: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511619427.003>
- Ariel, Mira. 1998. Three grammaticalization paths for the development of person verbal agreement in Hebrew. In Jean-Pierre Koenig (ed.), *Discourse and cognition: Bridging the gap*, 93–111. Stanford: CSLI/Cambridge University Press.
- Ariel, Mira. 2004. Accessibility marking: Discourse functions, discourse profiles, and processing cues. *Discourse Processes* 37(2). 91–116. DOI: [https://doi.org/10.1207/s15326950dp3702\\_2](https://doi.org/10.1207/s15326950dp3702_2)
- Armon-Lotem, Sharon. 2008. Subject use and the acquisition of verbal agreement in Hebrew. In Natalia Gagarina & Insa Gülzow (eds.), *The acquisition of verbs and their grammar: The effect of particular languages*, 45–67. Dordrecht: Springer. DOI: [https://doi.org/10.1007/1-4020-4335-X\\_2](https://doi.org/10.1007/1-4020-4335-X_2)

- Arnon, Inbal. 2010. Rethinking child difficulty: The effect of NP type on children's processing of relative clauses in Hebrew. *Journal of Child Language* 37(1). 27–57. DOI: <https://doi.org/10.1017/S030500090900943X>
- Baayen, R. Harald & Michael Ramscar. 2015. Abstraction, storage and naive discriminative learning. In Ewa Dabrowska & Dagmar Divjak (eds.), *Handbook of Cognitive Linguistics*, 100–120. Berlin, Boston: De Gruyter Mouton.
- Bendavid, Aba. 1988. *Biblical and Mishnaic Hebrew*. Tel Aviv: Dvir.
- Berman, Ruth A. 1978. *Modern Hebrew structure*. Tel-Aviv: Tel-Aviv University Pub. Projects.
- Berman, Ruth A. 1980. The case of an (s)VO language: Subjectless constructions in modern Hebrew. *Language* 56(4). 759–776. DOI: <https://doi.org/10.2307/413487>
- Berman, Ruth A. 1985. *Acquisition of Hebrew*. Hillsdale, NJ: Lawrence Erlbaum.
- Berman, Ruth A. 1990. On acquiring an (S)VO language: Subjectless sentences in children's Hebrew. *Linguistics* 28(6). 1135–1166. DOI: <https://doi.org/10.1515/ling.1990.28.6.1135>
- Berman, Ruth A. & Dan I. Slobin. 1994. *Relating events in narrative: A crosslinguistic developmental study*. Hillsdale, NJ: Lawrence Erlbaum.
- Bidgood, Amy, Ben Ambridge, Julian M. Pine & Caroline F. Rowland. 2014. The retreat from locative overgeneralisation errors: A novel verb grammaticality judgment study. *PLOS one* 9(5). DOI: <https://doi.org/10.1371/journal.pone.0097634>
- Bloom, Lois. 1993. *The transition from infancy to language: Acquiring the power of expression*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511752797>
- Bloom, Paul. 1990. Syntactic distinctions in child language. *Journal of Child Language* 17(2). 343–355. DOI: <https://doi.org/10.1017/S0305000900013805>
- Blum-Kulka, Shoshana, Michal Hamo & Talia Habib. 2010. Explanations in naturally occurring peer talk: Conversational emergence and function, thematic scope, and contribution to the development of discursive skills. *First Language* 30(3–4). 440–460. DOI: <https://doi.org/10.1177/0142723710370528>
- Borer, Hagit. 1989. Anaphoric agr. In Osvaldo Jaeggli & Kenneth J. Safir (eds.), *The null subject parameter*, 69–109. Dordrecht: Kluwer. DOI: [https://doi.org/10.1007/978-94-009-2540-3\\_3](https://doi.org/10.1007/978-94-009-2540-3_3)
- Borer, Hagit. 1995. The ups and downs of Hebrew verb movement. *Natural Language & Linguistic Theory* 13(3). 527–606. DOI: <https://doi.org/10.1007/BF00992740>
- Borer, Hagit & Kenneth Wexler. 1987. The maturation of syntax. In Thomas Roeper & Edwin Williams (eds.), *Parameter setting*, 123–172. Dordrecht: Springer. DOI: [https://doi.org/10.1007/978-94-009-3727-7\\_6](https://doi.org/10.1007/978-94-009-3727-7_6)
- Brener, Rosemary. 1983. Learning the deictic meaning of third person pronouns. *Journal of Psycholinguistic Research* 12(3). 235–262.
- Brooks, Patricia J. & Michael Tomasello. 1999. How children constrain their argument structure constructions. *Language* 75(4). 720–738. DOI: <https://doi.org/10.2307/417731>
- Brown, Roger. 1973. *A first language: The early stages*. Oxford: Harvard U. Press. DOI: <https://doi.org/10.4159/harvard.9780674732469>
- Bubenik, Vit. 1989. On the origins and elimination of ergativity in Indo-Aryan languages. *Canadian Journal of Linguistics/Revue canadienne de linguistique* 34(4). 377–398. DOI: <https://doi.org/10.1017/S0008413100024294>
- Chomsky, Noam. 2000. *New horizons in the study of language and mind*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511811937>
- Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.

- Clahsen, Harald. 1990. Constraints on parameter setting: A grammatical analysis of some acquisition in stages in German child language. *Language Acquisition* 1(4). 361–391. DOI: [https://doi.org/10.1207/s15327817la0104\\_3](https://doi.org/10.1207/s15327817la0104_3)
- Comrie, Bernard. 1978. Ergativity. In Winfred Philipp Lehmann (ed.), *Syntactic typology: Studies in the phenomenology of language*, 329–394. Austin, TX: University of Texas Press.
- Croft, William. 2001. *Radical construction grammar: Syntactic theory in typological perspective*. Oxford, New York: Oxford University Press.
- Dasinger, Lisa & Cecile Toupin. 1994. The development of relative clause functions in narrative. In Ruth A. Berman & Dan A. Slobin (eds.), *Relating events in narrative: A crosslinguistic developmental study*, 239–284. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Dattner, Elitzur. 2008. *Lower transitivity constructions in Hebrew: The case of motion verbs*. Tel Aviv: Tel Aviv University MA dissertation.
- Dattner, Elitzur. 2015. *Mapping Hebrew dative constructions*. Tel Aviv: Tel Aviv University dissertation.
- Dattner, Elitzur. 2019. The Hebrew dative: Usage patterns as discourse profile constructions. *Linguistics* 57(5). 1073–1110. DOI: <https://doi.org/10.1515/ling-2019-0022>
- De Villiers, Jill. 1995. Empty categories and complex sentences: The case of wh-questions. In Paul Fletcher & Brian MacWhinney (eds.), *Handbook of child language*, 508–540. Oxford: Blackwell. DOI: <https://doi.org/10.1111/b.9780631203124.1996.00021.x>
- Divjak, Dagmar & Nick Fieller. 2014. Finding structure in linguistic data. In Dylan Glynn & Justyna A. Robinson (eds.), *Corpus methods for semantics: Quantitative studies in polysemy and synonymy*, 405–441. Amsterdam and Philadelphia: John Benjamins Publishing Company. DOI: <https://doi.org/10.1075/hcp.43.16div>
- Dixon, R. M. W. 1979. Ergativity. *Language* 55(1). 59–138. DOI: <https://doi.org/10.2307/412519>
- Dixon, R. M. W. 1994. *Ergativity*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511611896>
- Doron, Edit. 2005. *VSO and left-conjunct agreement: Biblical Hebrew vs Modern Hebrew*. Berlin: Mouton.
- Dromi, Esther & Ruth A. Berman. 1986. Language-specific and language general in developing syntax. *Journal of Child Language* 13(2). 371–387. DOI: <https://doi.org/10.1017/S0305000900008114>
- Dryer, Matthew S. 1997. On the six-way word order typology. *Studies in Language* 21(1). 69–103. DOI: <https://doi.org/10.1075/sl.21.1.04dry>
- Du Bois, John W. 2007. The Stance Triangle. In Robert Englebretson (ed.), *Stancetaking in discourse: Subjectivity, evaluation, interaction*, 139–182. Amsterdam: John Benjamins Publishing Company. DOI: <https://doi.org/10.1075/pbns.164.07du>
- Du Bois, John W. 2017. Ergativity in discourse and grammar. In Jessica Coon, Diane Massam & Lisa deMena Travis (eds.), *The Oxford handbook of ergativity*, 23–58. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/oxfordhb/9780198739371.013.2>
- Du Bois, John W., Lorraine E. Kumpf & William J. Ashby. 2003. Introduction. In John W. Du Bois, Lorraine E. Kumpf & William J. Ashby (eds.), *Preferred argument structure: Grammar as architecture for function*, 1–10. Amsterdam and Philadelphia: John Benjamins Publishing Company. DOI: <https://doi.org/10.1075/sidag.14.03dub>
- Falk, Yehuda N. 2006. *Subjects and universal grammar: An explanatory theory*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511486265>
- Ferdinand, Astrid. 1996. *The development of functional categories: The acquisition of the subject in French*. The Hague: Holland Academic Graphics.
- Fischer, Kerstin. 2015. Conversation, construction grammar, and cognition. *Language and Cognition* 7(4). 563–588. DOI: <https://doi.org/10.1017/langcog.2015.23>

- Fletcher, Paul & Richard Ingham. 1995. *Grammatical impairment*. Cambridge: Blackwell.
- Forrester, Michael A. 1988. Young children's polyadic conversation monitoring skills. *First Language* 8(24). 201–226. DOI: <https://doi.org/10.1177/014272378800802401>
- Forrester, Michael A. & Sarah Cherrington. 2009. The development of otherrelated conversational skills: A case study of conversational repair during the early years. *First Language* 29(2). 166–191. DOI: <https://doi.org/10.1177/0142723708094452>
- Fox, Barbara A. & Sandra A. Thompson. 1990. A discourse explanation of the grammar of relative clauses in English conversation. *Language* 66(2). 297–316. DOI: <https://doi.org/10.2307/414888>
- Friedmann, Naama & Rama Novogrodsky. 2004. The acquisition of relative clause comprehension in Hebrew: A study of SLI and normal development. *Journal of Child Language* 31(3). 661–681. DOI: <https://doi.org/10.1017/S0305000904006269>
- Frizelle, Pauline, Jennifer Harte, Paul Fletcher & Fiona Gibbon. 2018. Investigating the effect of regional native accents on sentence comprehension in children with language impairment. *International Journal of Speech-Language Pathology* 20(4). 434–446. DOI: <https://doi.org/10.1080/17549507.2017.1293734>
- Gallaway, Clare & Brian J. Richards. 1994. *Input and interaction in language acquisition*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511620690>
- Givón, Talmy. 1977. The drift from VSO to SVO in Biblical Hebrew: The pragmatics of tense-aspect. In Charles N. Li (ed.), *Mechanisms of syntactic change*, 181–254. Austin, TX: University of Texas Press.
- Gleitman, Lila R. & Jane Gillette. 1995. The role of syntax in verb learning. In Paul Fletcher & Brian MacWhinney (eds.), *The handbook of child language*, 413–427. Cambridge, MA: Blackwell. DOI: <https://doi.org/10.1111/b.9780631203124.1996.00017.x>
- Goldberg, Adele E. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago, IL: University Of Chicago Press.
- Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
- Goldenberg, Gideon. 1998. *Studies in Semitic linguistics; selected writings*. Jerusalem: Magnes Press, Hebrew University.
- Goodwin, Marjorie & Amy Kyratzis. 2012. Peer language socialization. In Alessandro Duranti, Elinor Ochs & Bambi B. Schieffelin (eds.), *The handbook of language socialization*, 391–419. Oxford: Blackwell. DOI: <https://doi.org/10.1002/9781444342901.ch16>
- Grinstead, John Allen Ray. 1998. *Subjects, sentential negation and imperatives in child Spanish and Catalan*. Los Angeles, CA: University of California Los Angeles Doctoral dissertation.
- Grinstead, John Allen Ray. 2000. Constraints on the computational component vs. grammar in the lexicon: A discussion of Bates & Goodman. *Journal of Child Language* 27(3). 737–743. DOI: <https://doi.org/10.1017/S0305000900004335>
- Guerriero, Sonia, Amy Cooper, Yuriko Oshima-Takane & Yoko Kuriyama. 2001. A discourse-pragmatic explanation for argument realization and omission in English and Japanese children's speech. In Anna H.-J. Do, Laura Domínguez & Aimee Johansen (eds.), *Proceedings of the 25th annual Boston University Conference on Language Development*, 319–330. Somerville, MA: Cascadilla Press.
- Hacohen, Aviya & Jeannette Schaeffer. 2007. Subject realization in early Hebrew/English bilingual acquisition: The role of crosslinguistic influence. *Bilingualism: Language and Cognition* 10(3). 333–344. DOI: <https://doi.org/10.1017/S1366728907003100>



- Haendler, Yair & Flavia Adani. 2018. Testing the effect of an arbitrary subject pronoun on relative clause comprehension: A study with Hebrew speaking children. *Journal of Child Language* 45(4). 959–980. DOI: <https://doi.org/10.1017/S0305000917000599>
- Haspelmath, Martin & Andrea Sims. 2010. *Understanding morphology*. London: Routledge.
- Hatav, Galia. 2000. (Free) Direct discourse in Biblical Hebrew. *Hebrew Studies* 41. 7–30. DOI: <https://doi.org/10.1353/hbr.2000.0063>
- Helasvuo, Marja-Liisa & Aki-Juhani Kyröläinen. 2016. Choosing between zero and pronominal subject: Modeling subject expression in the 1<sup>st</sup> person singular in Finnish conversation. *Corpus Linguistics and Linguistic Theory* 12(2). 263–299. DOI: <https://doi.org/10.1515/cllt-2015-0066>
- Hoff, Erika. 2010. Context effects on young children’s language use: The influence of conversational setting and partner. *First Language* 30(3–4). 461–472. DOI: <https://doi.org/10.1177/0142723710370525>
- Howard, Alice Ann, Lara Mayeux & Letitia R. Naigles. 2008. Conversational correlates of children’s acquisition of mental verbs and a theory of mind. *First Language* 28(4). 375–402. DOI: <https://doi.org/10.1177/0142723708091044>
- Husson, François, Sébastien Lê & Jérôme Pagès. 2017. *Exploratory multivariate analysis by example using R*. Abingdon: Chapman and Hall/CRC.
- Hyams, Nina & Kenneth Wexler. 1993. On the grammatical basis of null subjects in child language. *Linguistic Inquiry* 24(3). 421–459.
- Keenan, Edward L. & Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8(1). 63–99.
- Kirjavainen, Minna, Evan Kidd & Elena Lieven. 2017. How do language-specific characteristics affect the acquisition of different relative clause types? Evidence from Finnish. *Journal of Child Language* 44(1). 120–157. DOI: <https://doi.org/10.1017/S0305000915000768>
- Köymen, Bahar & Amy Kyratzis. 2014. Dialogic syntax and complement constructions in toddlers’ peer interactions. *Cognitive Linguistics* 25(3). 497–521. DOI: <https://doi.org/10.1515/cog-2014-0028>
- Kuzar, Ron. 2012. *Sentence patterns in English and Hebrew*. Amsterdam and Philadelphia: John Benjamins Publishing Company. DOI: <https://doi.org/10.1075/cal.12>
- Lambrecht, Knud. 1994. *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511620607>
- Lê, Sébastien, Julie Josse & François Husson. 2008. FactoMineR: An R package for multivariate analysis. *Journal of Statistical Software* 25(1). 1–18. DOI: <https://doi.org/10.18637/jss.v025.i01>
- Levy, Yonata & Anne Vainikka. 2000. The development of a mixed null subject system: A cross-linguistic perspective with data on the acquisition of Hebrew. *Language Acquisition* 8(4). 363–384. DOI: [https://doi.org/10.1207/S15327817LA0804\\_03](https://doi.org/10.1207/S15327817LA0804_03)
- Li, Charles N. & Sandra A. Thompson. 1976. Subject and topic: A new typology of language. In Charles N. Li (ed.), *Subject and topic*, 457–489. New York: Academic Press.
- Marantz, Alec. 1982. Re reduplication. *Linguistic Inquiry* 13(3). 435–482.
- Matthei, Edward H. 1987. Subject and agent in emerging grammars: Evidence for a change in children’s biases. *Journal of Child Language* 14(2). 295–308. DOI: <https://doi.org/10.1017/S0305000900012940>
- McClure, Kathleen, Julian M. Pine & Elena V.M. Lieven. 2006. Investigating the abstractness of children’s early knowledge of argument structure. *Journal of Child Language* 33(4). 693–720. DOI: <https://doi.org/10.1017/S0305000906007525>

- Melnik, Nurit. 2006. A constructional approach to verb-initial constructions in Modern Hebrew. *Cognitive Linguistics* 17(2). 153–198. DOI: <https://doi.org/10.1515/COG.2006.004>
- Melnik, Nurit. 2007. Extending partial pro-drop in Modern Hebrew: A comprehensive analysis. In Stefan Müller (ed.), *Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar, Stanford Department of Linguistics and CSLI's LinGO Lab*, 173–193. Stanford, CA: CSLI Publications.
- Murase, Toshiki, Philip S. Dale, Tamiko Ogura, Yukie Yamashita & Aki Mahieu. 2005. Mother-child conversation during joint picture book reading in Japan and the USA. *First Language* 25(2). 197–218. DOI: <https://doi.org/10.1177/0142723705050899>
- Oshima-Takane, Yuriko. 1988. Children learn from speech not addressed to them: The case of personal pronouns. *Journal of Child Language* 15(1). 95–108. DOI: <https://doi.org/10.1017/S0305000900012071>
- Paradis, Johanne & Samuel Navarro. 2003. Subject realization and crosslinguistic interference in the bilingual acquisition of Spanish and English: What is the role of the input? *Journal of Child Language* 30(2). 371–393. DOI: <https://doi.org/10.1017/S0305000903005609>
- Perek, Florent. 2015. *Argument structure in usage-based construction grammar: Experimental and corpus-based perspectives*. Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/cal.17>
- Preminger, Omer. 2009. Failure to agree is not a failure:  $\varphi$ -agreement with post-verbal subjects in Hebrew. *Linguistic Variation Yearbook* 9(1). 241–278. DOI: <https://doi.org/10.1075/livy.9.07pre>
- R Core team. 2018. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Ramscar, Michael, Daniel Yarlett, Melody Dye, Katie Denny & Kirsten Thorpe. 2010. The effects of feature-label-order and their implications for symbolic learning. *Cognitive Science* 34. 909–957. DOI: <https://doi.org/10.1111/j.1551-6709.2009.01092.x>
- Ravid, Dorit. 1995. *Language change in child and adult Hebrew: A psycholinguistic perspective*. Oxford: Oxford University Press.
- Ravid, Dorit & Dalia Cahana-Amitay. 2005. Verbal and nominal expressions in narrating conflict situations in Hebrew. *Journal of Pragmatics* 37(2). 157–183. DOI: <https://doi.org/10.1016/j.pragma.2004.08.006>
- Romberg, Alexa R. & Jenny R. Saffran. 2010. Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science* 1(6). 906–914. DOI: <https://doi.org/10.1002/wcs.78>
- Salazar Orvig, Anne, Haydée Marcos, Aliyah Morgenstern, Rouba Hassan, Jocelyne Leber-Marin & Jacques Parès. 2010. Dialogical factors in toddlers' use of clitic pronouns. *First Language* 30(3–4). 375–402. DOI: <https://doi.org/10.1177/0142723710379957>
- Santelmann, Lynn, Stephanie Berk, Jennifer Austin, Shamitha Somashekar & Barbara Lust. 2002. Continuity and development in the acquisition of inversion in yes/no questions: Dissociating movement and inflection. *Journal of Child Language* 29(4). 813–842. DOI: <https://doi.org/10.1017/S0305000902005299>
- Schachter, Paul. 1976. The subject in Philippine languages: Topic, actor, actor-topic, or none of the above. In Charles N. Li (ed.), *Subject and topic*, 491–518. New York: Academic Press.
- Schuele, C. Melanie. 2010. The many things language sample analysis has taught me. *Perspectives on Language Learning and Education* 17(1). 32–37. DOI: <https://doi.org/10.1044/ll17.1.32>

- Schwarzwald, Ora R. 2002. *Modern Hebrew morphology*. Tel Aviv: Open University.
- Serratrice, Ludovica. 2005. The role of discourse pragmatics in the acquisition of subjects in Italian. *Applied psycholinguistics* 26(3). 437–462. DOI: <https://doi.org/10.1017/S0142716405050241>
- Serratrice, Ludovica & Shanley E. Allen. 2015. *The acquisition of reference*. Amsterdam and Philadelphia: John Benjamins Publishing Company.
- Shin, Naomi Lapidus. 2016. Children's Spanish subject pronoun expression: A developmental change. In Sandro Sessarego & Fernando Tejedo-Herrero (eds.), *Spanish language and sociolinguistic analysis*, 155–176. Amsterdam and Philadelphia: John Benjamins Publishing Company.
- Shlonsky, Ur. 1997. *Clause structure and word order in Hebrew and Arabic: An essay in comparative Semitic syntax*. Oxford: Oxford University Press.
- Snow, Catherine E. & Charles A. Ferguson. 1977. *Talking to children*. Cambridge: Cambridge University Press.
- Tomasello, Michael. 1992. *First verbs: A case study of early grammatical development*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511527678>
- Tomasello, Michael. 2003. *Constructing a language: A usage-based approach to child language acquisition*. Cambridge: Cambridge University Press.
- Tomasello, Michael. 2009. The usage-based theory of language acquisition. In Edith L. Bavin (ed.), *The Cambridge handbook of child language*, 69–87. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511576164.005>
- Valian, Virginia. 1990. Null subjects: A problem for parameter-setting models of language acquisition. *Cognition* 35(2). 105–122. DOI: [https://doi.org/10.1016/0010-0277\(90\)90011-8](https://doi.org/10.1016/0010-0277(90)90011-8)
- Valian, Virginia. 1991. Syntactic subjects in the early speech of American and Italian children. *Cognition* 40(1–2). 21–81. DOI: [https://doi.org/10.1016/0010-0277\(91\)90046-7](https://doi.org/10.1016/0010-0277(91)90046-7)
- Van Valin, Robert D. Jr. 2004. Semantic macroroles in role and reference grammar. In Rolf Kailuweit (ed.), *Semantische rollen*, 62–82. Narr: Tübingen.
- Van Valin, Robert D. Jr. & Randy J. LaPolla. 1997. *Syntax: Structure, meaning, and function*. Cambridge, MA: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9781139166799>
- Veneziano, Edy. 2010. Conversation in language development and use: An introduction. *First Language* 30(3–4). 241–249. DOI: <https://doi.org/10.1177/0142723710380531>
- Veneziano, Edy. 2018. Learning conversational skills and learning from conversation. In Amalia Bar-On & Dorit Ravid (eds.), *Handbook of communications disorders: Theoretical, empirical, and applied linguistic perspectives*, 311–328. Berlin, Boston: De Gruyter Mouton. DOI: <https://doi.org/10.1515/9781614514909-016>
- Veneziano, Edy & Eve V. Clark. 2016. Early verb constructions in French: Adjacency on the left edge. *Journal of Child Language* 43(6). 1193–1230. DOI: <https://doi.org/10.1017/S0305000915000471>
- Wittgenstein, Ludwig. 1953. *Philosophical investigations*. New York: McMillan.
- Yelland, Phillip M. 2010. An introduction to correspondence analysis. *The Mathematica Journal* 12(1). 86–109. DOI: <https://doi.org/10.3888/tmj.12-4>
- Ziv, Yael. 1988. Word order in children's literature: FSP and markedness. In Yishay Tobin (ed.), *The Prague school and its legacy*, 123–144. Amsterdam and Philadelphia: John Benjamins Publishing Company. DOI: <https://doi.org/10.1075/llsee.27.12ziv>
- Zwiling, Racheli. 2009. *Noun plurals in children's peer talk 2–8 years*. Tel Aviv: Tel Aviv University MA dissertation.

**How to cite this article:** Dattner, Elitzur, Liron Kertes, Racehli Zwillig and Dorit Ravid. 2019. Usage patterns in the development of Hebrew grammatical subjects. *Glossa: a journal of general linguistics* 4(1): 129.1–28. DOI: <https://doi.org/10.5334/gjgl.928>

**Submitted:** 25 February 2019    **Accepted:** 29 August 2019    **Published:** 19 November 2019

**Copyright:** © 2019 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.



*Glossa: a journal of general linguistics* is a peer-reviewed open access journal published by Ubiqity Press.

OPEN ACCESS The Open Access icon, which is a stylized 'a' inside a circle.