

# The Autonomous Contribution of Syntax and Pragmatics to the Acquisition of the Hebrew Definite Article<sup>\*</sup>

Idit Avram and Sharon Armon-Lotem  
Bar Ilan University

## 1. Introduction

This paper focuses on the acquisition of syntactic and pragmatic aspects of the definite system in Hebrew, in comparison to English, in connection to the Theory of Mind. Definiteness is a grammatical marker of the coordination and differentiation relation between a speaker, a hearer, and their knowledge about a referent. In other words, definiteness is concerned with the grammaticalization of identifiability and nonidentifiability of referents on the part of a speaker or hearer. Thus, the use of definiteness depends on acquiring syntactic, semantic, and pragmatic knowledge. The present study investigates the autonomous contribution of syntax and pragmatic to the acquisition of the Hebrew definite system.

Different languages mark definiteness in typologically different ways. In English all nouns require an article, definite or indefinite, preceding the whole noun-phrase. In Hebrew, definiteness is marked by a definite prefix on every definite noun and any adjective modifying it. In both English and Hebrew, unlike German, for example, the definite marker does not agree in gender or number with the head noun.

### *1.1 Definiteness in English*

Syntactically, every singular count noun in English requires an indefinite or definite article (e.g., *\*(a/the) boy*). Plural count nouns and mass nouns require the definite article whenever pragmatically appropriate (e.g., *(the) boys*, *(the) sugar*). The

---

\* An extended version of this paper is to appear in Anne-Marie Di Sciullo (ed.) Universal Grammar and the External systems: Essays on Language, Brain and Computation. John Benjamins Publishing Company.

choice of articles depends on the knowledge of the speaker of her own knowledge and of others' knowledge as well.

Schaeffer (1997) discusses the properties of the definite and indefinite articles in adult English. The referential definite 'the' is used when the referent is known to both speaker and hearer, as in (1):

- 1) There is a car and a ball on the carpet. The car is red.

The object in the world that 'the car' refers to is known by both speaker and hearer, due to the fact that it was mentioned in the previous discourse. The referent can also be part of world knowledge shared by both speaker and hearer, as in (2):

- 2) The sun is out.

There is only one object in the world that is known as 'the sun'.

The referential indefinite 'a' is used when the referent is known only to the speaker, as in (3):

- 3) I bought a book yesterday.

The indefinite nominal 'book' does have a referent, but only the speaker knows it.

The non-referential indefinite 'a' is used when the referent is unknown to both the speaker and hearer, as in (4):

- 4) I haven't seen a movie for ages.

There is no movie in the world that could be the referent for 'a movie'. Schaeffer concluded from the above that in adult English the form of the definite article is based on definiteness rather than referentiality, since the latter relies solely on speaker's knowledge, while in fact both speaker's and hearer's knowledge are crucial for definiteness.

Schaeffer proposes that definiteness and referentiality should be syntactically marked by features such as [speaker] and [hearer] on the D-head of the DP, as specified in (5):

- (5) a. A referential definite nominal expression is marked with the features [speaker] and [hearer].
- b. A referential indefinite nominal expression is marked with the feature [speaker].
- c. A non-referential nominal expression is not marked with these features.

Schaeffer further claims that non-referential indefinite nominal expressions are NumPs, without a D-head, and are not marked with [speaker] or [hearer] features. Consequently, the pragmatic distinction between speaker and hearer is crucial to the acquisition of the properties of D.

### 1.2 Definiteness in Hebrew

Definiteness marking in Hebrew is less uniform than in English or German, but resembles other Semitic languages, such as Arabic (Wintner, 2000). The only definite marker in Hebrew, 'ha-', is a nominal prefix that does not inflect, as in (6), rather than a function word marking a whole phrase. It is used in agreement both on nouns and on the adjectives modifying them, as shown in (7). In the direct object position, definite NPs are introduced by the definite accusative case marker 'et' which has the characteristics of a preposition, as shown in (8):

- 6) a. Indefinite NP: *yeled shata mic*  
 boy drank juice  
 'A boy drank some juice'
- b. Definite NP: *raiti yeled. ha-yeled shata mic*  
 I-saw boy. the-boy drank juice  
 'I saw a boy. The boy drank some juice'
- 7) Agreement: *ha-yeled ha-gadol shata mic*  
 The-boy the-big drank juice  
 'The big boy drank some juice'

- 8) Direct Object: *kaniti mic. yeled shata \*(et) ha-mic*  
 I-bought juice. Boy drank acc the-juice  
 'I bought some juice. A boy drank the juice'

Non-discourse-related definite (generic) nouns in Hebrew, unlike English, may appear without the definite article 'ha', depending on their syntactic position and semantic reading, as in (9):

- 9) a. *Yesh (\*ha-)shemesh ba-xuc*  
 cop(exist) (\*the-)sun outside  
 'It's sunny outside' (literally: the sun exists outside)
- b. *Hine (ha-)shemesh*  
 Here (the-)sun  
 'Here is the sun'
- c. *Ani roe (et ha-)shemesh*  
 I see (acc the-)sun  
 'I see the sun'
- d. *Ani mistakelet al \*(ha-)shemesh*  
 I look at (the-)sun  
 'I look at the sun.'

In (9a), with the existential copula *yesh*, the definite article is ungrammatical. Sentences (b) and (c) have an existential reading too, reflected in the possibility of replacing the optional definite article *the* existential copula *yesh*. Thus, the article is optional. In sentence (d), the article is syntactically obligatory since it follows a preposition, which takes a DP as its complement. These characteristics of the definite system in Hebrew make it less uniform and more sensitive to a variety of syntactic and semantic factors.

## 2. The Acquisition of Definiteness

The acquisition of the definiteness system poses a hard and complex communicative problem to the child (Zur, 1983). The proper use of the definiteness system is a combination of a syntactic task and a pragmatic task, which requires an

ability to consider the hearer's knowledge, overriding the speaker's egocentric tendency in order to create a coherent referent system for the hearer. The speaker needs to know that her knowledge is different from the hearer's own knowledge, in order to share her knowledge with the hearer. This means that the child has to distinguish between objects that exist in reality and those which exist only in her mind. Children seem to lack the ability to differentiate speaker from hearer at the first stages of language acquisition, and have to compensate for this deficit by other means.

### *2.1 Acquisition of Definiteness in English*

Schaeffer (1997) conducted a study on the acquisition of definiteness by English-speaking children. She argues that while the use of articles is close to adult-like by the age of four, at 2 and 3, children make some syntactic errors. She claims that children initially rely on referentiality in their choice of articles, since referentiality is a semantic notion that does not rely on hearer's knowledge, but solely on speaker's knowledge. Thus, the syntactic errors result from the lack of a pragmatic principle, 'the concept of non-shared knowledge', which differentiates the knowledge of the speaker and the hearer. She further claims that development within the pragmatic component will lead to a development within syntax. More specifically, she argues that non-referential indefinite nominal expressions are NumPs, without a D-head, and are not marked with [speaker] or [hearer] features. Consequently, the pragmatic distinction between speaker and hearer is crucial to the acquisition of the properties of D.

Schaeffer (1997) predicts that children will overgenerate the definite referential article '*the*' to contexts which require the indefinite referential *a*, indicating lack of pragmatic knowledge about the speaker/hearer differentiation. On the other hand, children will drop the definite article '*the*' in definite contexts, and the indefinite

article 'a' in referential indefinite contexts due to lack of syntactic knowledge about the use of D. She found that English-speaking 2-year-olds, and to a lesser extent 3-year-olds drop articles in referential contexts and overgenerate definite articles in indefinite contexts. Schaeffer suggests that these errors result from the child's lack of the concept of non-shared knowledge, a pragmatic concept which indicates that speaker's knowledge and hearer's knowledge are always independent. Once this concept has been acquired, the child distinguishes the two features and marks them on D.

### *2.2 Acquisition of Definiteness in Hebrew*

Zur (1983), found that just as in English, the definite marker is already used by the age of two. Initially, up to the age of 3;6, Zur found many errors in the use of definiteness. At the age of three to five there was an overgeneralization of the rules, and lack of knowledge of the constraints limiting the application of the rules. The age of five was found to be critical for the acquisition of the definite system.

Syntactically, as in other languages, children start by using bare nouns, which are extended within the earliest word combinations into full NPs, where quantifiers and determiners are used in [Spec, NP] (Armon-Lotem, 1996). Since there is not enough space to host more than one modifier, the projection extended into NumP and DP. This syntactic structure makes it possible for the child to extend the number of possible features which are identified. The productive use of the definite article *ha-* is a partial proof of the above. Both Zur's and Armon-Lotem's findings were not discussed in light of the distinctions between speaker and hearer knowledge, that is the concept of non-shared knowledge.

### *2.3 The Concept of Non-Shared Knowledge and the Theory of Mind*

The concept on non-shared knowledge (Schaeffer, 1997) has been associated with some aspects of 'The Theory of Mind' (TOM). Malle claims that theory of mind

“refers to the ability to represent, conceptualize, and reason about mental states” (2001:3) of self and others. Piaget (1959) was the first to investigate children’s understanding of the mind, and conclude that children are egocentric and make no attempt to understand the other’s point of view. If the speaker, being a young child, cannot read the mind of the other, she automatically attributes her own knowledge to the other, ignoring the distinction between the knowledge of the two minds. Within TOM, conceptual understanding of desire (age 2) is followed by the conceptual understanding of belief (age 3), and false belief (age 4). Understanding the desires and beliefs of the other is understanding that speaker and hearer do not share the same knowledge.

### **3. Methodology**

The current study tests Schaeffer’s (1997) predictions, in a typologically different language, Hebrew. The aim of the study is to see whether Hebrew, despite the typological differences in structure (e.g., prefix vs. function word) and acquisition (i.e., definiteness is mastered later in Hebrew than in English), will show the same discrepancy between syntax and pragmatics as English does. Furthermore, this study tests whether the findings for Hebrew lend themselves to an analysis that ties the pragmatic aspects of definiteness to the Theory of Mind.

The following predictions emerge from Schaeffer’s findings (1997), taking into account typological differences:

1. Children will drop the definite article ‘ha-‘ in referential definite contexts, due to limited syntactic knowledge.
2. Children will overgenerate the definite article ‘ha-‘ in contexts which require the indefinite referential article (in Hebrew- zero article), until they acquire the concept of non-shared knowledge.

3. Children will drop the definite article for non-discourse-related definite NPs, until they acquire the relevant world knowledge

### *3.1 Subjects*

32 Hebrew-speaking children aged 2 – 5, divided into three groups, and an adult control group were tested. The children, all from middle socioeconomic status (SES), attended different preschools across Israel. Most interviews took place at the preschool in a quiet room with one or two experimenters. Some of the sessions, however, took place at the subject's home. All sessions were tape-recorded and later on transcribed.

### *3.2 Method*

To test the above predictions, a toy elicitation task (Crain & Thornton 1998; Thornton 1996) was used. The experimenter manipulated the toy props while a blindfolded puppet was listening to the story. Being blindfolded, the puppet needed the child's help to understand the story. Since the puppet was blindfolded, the child could not use deictics like 'this' to identify the objects. Therefore, the only way the child could correctly identify the objects for the puppet was by naming them.

The experiment consisted of up to eighteen scenarios, at least three in each of the following categories:

- a) The referential definite existent
- b) The non-discourse-related definite NP.
- c) The referential indefinite existent
- d) The referential indefinite non-existent.
- e) The non-referential indefinite.

All the scenarios were piloted with an adult control group. Scenarios were presented in a random order, over two to three sessions. The first session started with up to six



‘warm-up’ stories containing a personal pronoun, in order to verify that the children understand the task. Warm-up stories were used in the following sessions only when necessary. (10) is an example of the task in the definite – referential context:

- 10) Situation: Mr. Bunny is blindfolded.  
 There is a ball and a car on the table.  
 Mr. Lion enters.
- Mr. Bunny: Who is it  
 Mr. Lion: It’s me Mr. Lion  
 Mr. Bunny: What’s on the table  
 Mr. Lion: A ball and a car. I have an idea
- Situation: Mr. Lion pushes the car.  
 Mr. Bunny: What did Mr. Lion do?
- Child: *hu daxaf et ha-oto* / ‘He pushed the car’  
 \* *hu daxaf oto* / ‘\* He pushed a car’

### 3.3 Categories of Analysis

The five aforementioned contexts were analyzed into four categories, collapsing together the two indefinite referential contexts since the results were identical. The four categories were:

- a) Definite discourse-related contexts.
- b) Definite non-discourse-related contexts.
- c) Indefinite referential contexts.
- d) Indefinite non-referential contexts.

Data in each category were analyzed, per age group, regarding the number of errors and type of errors, the omission or addition of the definite article.

## 4. Findings

Children of all age groups used the definite article, with a very few errors. There were hardly any additions of the definite article in the indefinite context, but there were some omissions of the definite articles in both obligatory contexts. Analyzing

the types of errors, it was found that the addition of the definite article was limited to the twos (M 2;3). It occurred in 10% (5/54) of the referential indefinite contexts, but never in the non-referential contexts. Dropping of the definite article was found in 13% (9/70) of discourse-related definite context, up to age 4 (M 4;6), with the twos showing the lowest error rate (1/15) and the threes showing the highest error rate (18% - 5/28). Dropping of the definite article was found in 21% (21/100) of non-discourse-related definite contexts, up to age 4 (M 4;6), the fours' error rate being as high as the twos'. This is shown in Figure 1:

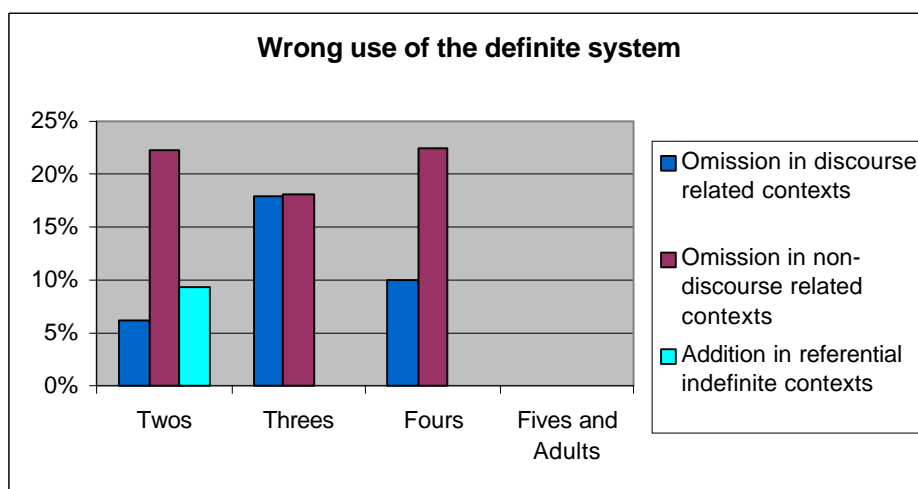


Figure 1– Error rate across categories (by age)

In sum, omission of definite articles is more frequent than addition of the article, and remains a problem for a longer period. The use of definiteness in non-discourse-related contexts seems to pose a consistent problem as late as the age of four, whereas the use of definiteness in referential indefinite contexts is problematic only for the twos.

## 5. Discussion

The Hebrew findings are consistent with Schaeffer's findings for English as far as additions are concerned. In both languages, two-years-old children overgeneralize the

definite article to indefinite referential contexts. It looks as if they used the definite article to mark referentiality rather than definiteness. From a syntactic perspective, it seems that they do not distinguish between speaker and hearer features, since they do not distinguish speaker and hearer knowledge. Since they never overgeneralize in the non-referential context, it seems safe to assume that in this context no feature is available and D is not used at all. Around the age of three they learn the pragmatic concept of non-shared knowledge and start using the definite article only in definite referential contexts.

In English, however, the system stabilizes both pragmatically and syntactically by the age of three, whereas in Hebrew, omissions of the definite article are found as late as four, and the question is why. Schaeffer argues that omissions around the age of two are due to lack of syntactic knowledge which makes it possible to use an NumP rather than a DP. These omissions were bound to disappear once children acquire the concept of non-shared knowledge, which enables them to mark the features ‘hearer’ and ‘speaker’ on D.

The Hebrew-speaking children seem to have acquired the concept of non-shared knowledge, as is evident from the absence of overgeneralizations of the definite article beyond the age of three. The question is whether we can attribute lack of syntactic knowledge to the Hebrew-speaking four-year-olds. As noted earlier, most omissions occur in the non-discourse-related definite context, a context which is not homogeneous with respect to the use of the definite article and requires world knowledge, rather than unique syntactic knowledge. It seems that while Hebrew-speaking children clearly differentiate the speaker’s knowledge from the hearer’s knowledge before the age of three, they do not know, by the age of four, which beliefs can be attributed to the hearer based on shared knowledge, that is, they are lacking the pragmatic concept of shared knowledge.

Why is the type of pragmatic knowledge which is acquired by the age of three enough to stabilize the acquisition of definiteness in English, but not in Hebrew? The answer seems to lie within the domain of syntax, or more specifically in the syntactic uniformity of the definiteness paradigm. Definiteness in English is a uniform paradigm. Articles are obligatory and are used systematically in both definite and indefinite contexts. In Hebrew, however, definiteness is a non-uniform paradigm which has only a definite article, and even this article is optional in some of the

contexts where it is obligatory in English, e.g., in non-discourse-related definite context.

Being optional, the use of the definite article is more sensitive to a wider range of syntactic, semantic, and pragmatic factors. One possible semantic factor is that the English indefinite article *a* has a referential interpretation and a quantifier interpretation (Fodor and Sag, 1982), while indefinite bare nouns in Hebrew lack the quantifier interpretation (Y. Greenberg, p.c.). For example, while the use of the indefinite article in ‘*a sun*’ indicates that there is more than one sun, the bare noun *shemesh* (sun in Hebrew) does not have this reading. Therefore, in English, the use of a definite article with such nouns is semantically obligatory, in order to avoid the ‘one of many’ reading. In Hebrew, since the bare nouns do not have this reading, non-discourse-related definite nouns, like the *sun* or the *moon*, do not necessarily require a definite article and its use relates to particular syntactic structures (e.g., after a phrasal verb), and the different pragmatic aspects of definiteness. This complexity influences the use of the definite article in definite discourse-related contexts, causing a marginal instability. To sum, Hebrew-speaking children acquire the concept of non-shared knowledge by the age of 3, but master the system only when they acquire the concept of shared knowledge, i.e., when they learn about the other’s beliefs.

## **6. Conclusion**

The Hebrew findings support Schaeffer’s findings regarding the separate contribution of pragmatic principles and syntactic principles to the acquisition of definiteness, and add to them a second distinction between the contributions of different pragmatic principles to the acquisition of definiteness. These findings, while answering some questions, raise many others. Does this second distinction mirror the child’s cognitive development? Is the concept of non-shared knowledge a linguistic

manifestation of self-other differentiation? Is the concept of shared knowledge a linguistic manifestation of self-other coordination? Is shared knowledge related to the Theory of Mind? Can one assume that the acquisition of the concept of non-shared knowledge (before age 3) follows the conceptual understanding of desire (age 2), but precedes the conceptual understanding of belief (age 3), whereas understanding of belief (age 3) and false belief (age 4) seems to facilitate the acquisition of the concept of shared knowledge (age 4-5)? Further research is needed in order to answer these questions and find out how the pragmatics of definiteness is related to the Theory of Mind.

### **References**

- Armon-Lotem, S. "The Minimalist Child: Parameters and Functional Heads in the Acquisition of Hebrew." Ph.D. Thesis, Tel-Aviv University, 1996.
- Astington, Janet, W. The Child's Discovery of the Mind. New York: Harvard UP, 1993.
- Brentano, F. "The Distinction between Mental and Physical Phenomena." Realism and the Background of Phenomenology. Ed. R.M. Chisholm. New York: Free Press, 1960.
- Chesterman, Andrew. On Definiteness: A Study with Special Reference to English and Finnish. New York: Cambridge UP, 1991.
- Crain S. & R. Thornton. Investigations in Universal Grammar. MIT Press, 1998.
- Fodor, Janet, D., and Ivan A. Sag. "Referential and Quantificational Indefinites." Linguistics and Philosophy 5 (1982): 355-398.
- Hawkins, J. Definiteness and Indefiniteness: A Study in Reference and Grammaticality Prediction. London: Croom Helm, 1978.

- Hoekstra, Teun, and Nina Hyams. "The Syntax and Interpretation of Dropped Categories in Child Language: A Unified Account." The Proceedings of the Fourteenth West Coast Conference on Formal Linguistics. Ed. Camacho Jose, Lina Choueiri, and Maki Watanabe. United States: CSLI Publications, 1996. 123-136.
- Malle, Bertram, F. "The Relation between Language and Theory of Mind in Development and Evolution." Presented at the symposium on *The Evolution of language out of Pre-language*, Institute of Cognitive and Decision Sciences. University of Eugene: Oregon, 2001.
- Piaget, Jean. The Language and Thought of the Child. London: Routledge and Kegan Paul, 1959.
- Schaeffer, Jeannette. "Articles in English Child Language." Paper presented at the Linguistic Society of America, Los Angeles, January 1997.
- Thornton, Rosalind. "Elicited Production," in D. McDaniel, C. McKee and H. S. Cairns, eds., Methods for Assessing Children's Syntax., 77-101, Cambridge, Massachusetts: MIT Press, 1996.
- Wellman, H.M. The Child's Theory of Mind. Cambridge, Mass.: Bradford Books, MIT Press, 1990.
- Wimmer, H., and Perner, J. "Beliefs about Beliefs: Representation and Constraining Function of Wrong Beliefs in Young Children's Understanding of Deception." Cognition 13 (1983): 103-128.
- Wintner, Shuly. "Definiteness in the Hebrew Noun Phrase." Journal of Linguistics 36:2 (2000): 319- 363.
- Zur, Batia. "On the Acquisition of Definiteness by Hebrew-Speaking Children." MA Paper, Tel-Aviv University, 1983. (In Hebrew)