

Acquisition of Definiteness in Emirati Arabic

Dimitrios Ntelitheos, *Department of Linguistics*, E-mail: Dimitrios_N@uaeu.ac.ae

Ali Idrissi, *Department of Linguistics*, Othaiba Tamimi, *Department of Linguistics*,
U.A.E. University, Al-Ain, P.O. Box: 17555, U.A.E.

Abstract

Crosslinguistic studies of language development show that determiners are omitted in early stages. Possible reasons for these errors are the unavailability of the higher functional nominal layer and consequently the lack of a projection hosting articles and the fact that young children lack theory of mind in that they do not have the ability to represent, conceptualize, and reason about the mental states of others. Evaluating these two approaches we checked determiner omission patterns in the speech of six Emirati children aged 1;11 to 4;4 years. We found that determiners are productively used in early stages but children do omit definite determiners in contexts where they are obligatory in the target language. Therefore, while unavailability of the D-system is contradicted by the data, the omission of definite determiners must be the result of the underspecification of the D-system governed by pragmatic constraints.

1. INTRODUCTION

Crosslinguistic studies on the acquisition of referring expressions have shown that children frequently omit definite determiners in early stages of language acquisition. From a morphosyntactic point of view the omission of definite determiners (and indefinite determiners where available) has been tied to the presence or absence in child grammar of the appropriate functional projections that host these determiners. Thus, a number of proposals (see for example [1], [2], and [3]), associate the lack of functional material in child speech to the fact that child grammar develops in an incremental fashion, starting from hierarchically low lexical material, with higher functional layers added in later stages. These ideas are inherent in the “small clause” hypothesis, developed in [1], or the “truncation model”, put forward in [2].

Other researchers have proposed that all functional categories are operative at the first stages of morphosyntactic development ([4], [5]). However, certain features which are obligatorily assigned a value in adult language are left underspecified in child language ([5]). Hoekstra et al ([5]) propose that what is lacking in children sentences where the definite determiner is dropped is the grammatical encoding of specificity. Children are somehow tied to the “here and now”, i.e. lack of pragmatic knowledge or awareness introduces underspecification of functional projections in child grammar. However, this does not mean that children are not sensitive to the presence of the functional projection that hosts determiners, as comprehension and production studies have shown ([6], [7]).

The pragmatic aspect of the acquisition of definiteness has been explored in detail in later studies [8], [9], [10] and others). Schaeffer in [8] proposed that children lack the pragmatic concept of ‘non-shared assumptions’, i.e. the fact that speaker and hearer assumptions are independent. Gundel, Ntelitheos, and Kowalski [9], show that the order of acquisition of referring forms follows the hierarchical order of referring expressions in terms of ‘givenness’ i.e. the cognitive status of the referent for the addressee. Thus, pronouns are acquired early and indefinites last, but children at age 3;00 or earlier use referring forms appropriately.

To evaluate these proposals we examined the use of D-type elements and especially the use of the definiteness prefix marker (ə)l-, in the speech of six Emirati Arabic-speaking children. The study is based on transcriptions of recordings of spontaneous speech, spanning six months (ages 1;6-4;4). We found that children from the early stages use definiteness markers productively, which seems to indicate that the D-system is intact. In the early stages (1;6-2;00) children almost exclusively use pronouns, demonstrative pronouns and bare nominal forms. In later stages (2;7-4;4) the definite marker appears in the right contexts but its frequency is significantly lower than that in adult speech and children prefer bare nouns instead. Furthermore, children optionally omit the determiner in contexts in which it is obligatory in the target language (usually proper names or in generic nouns). Finally, in very few cases children use definite determiners in contexts which would be unexpected in adult language.

These results seem to indicate a tension between syntactic and pragmatic factors in the acquisition of definiteness in Emirati Arabic. Correct use of the definite marker, as well as the use of proper names, pronouns and demonstratives supports the ‘full competence hypothesis’ in language acquisition ([4], [5]), which assumes that children have the full set of projections from the start. On the other hand, the high rate of definite marker omission, as evidenced by the low frequency of the marker in child data indicates that the D-projection is underspecified and children have not completely mastered the use of the marker. Finally, the extremely low rate of definite marker use in adult indefinite contexts indicates that children are sensitive to the memory and attention state of the interlocutors of referring expressions.

The paper is organized as follows: In Section (2) we discuss the D-system in Arabic and show that the definite determiner appears very frequently in adult speech, presumably because of its productive use in a number of syntactically well-defined environments. In Section (3), we provide information about the participants and the collection and transcription of the longitudinal data that forms the empirical basis for our discussion as well as the results with respect to definiteness marking. Section (4) contains our discussion. Finally, Section (5) provides concluding remarks and issues that need further development.

2. EXPRESSION OF DEFINITENESS IN ARABIC

Emirati Arabic noun phrases are syntactically marked as definite while indefinite noun phrases are morphologically unmarked. However, morphosyntactic marking does not always represent accurately the pragmatic notions of definiteness and specificity ([10]: 18). Definite noun phrases are either marked with the definite determiner prefix (*al-*) or a phonetically determined variant; within the construct state by agreement to the definiteness feature of the possessor DP (2); or by a suffixed possessive pronoun (3) (see [10], [11], [12]):

1. *ʕa l-baħər*
on D-sea
‘on the sea’
2. *beezaat l-kuuli*
money D-worker
‘the worker’s money’
3. *beet.hum [beettum]*
house.their (ms.)
‘their house’

Indefinites sometimes implore the numeral *wa:ħid* ‘one’ to refer to an indefinite specific entity.[11]:114 translates it as ‘a certain’, as in the following example:

4. *ra:ħat ħagg wa:ħid mt^ʕawwaʕ*
went.3SG.FEM to one religious.man
‘She went to a learned religious man.’

In addition, *wa:ħid* seems to modify only human nouns ([10]: 20)

Abstract, mass, collective and generic nouns require the definite article (examples from [11]:159):

5. *is^ʕ-s^ʕidg ʔaxyar min il-kidħbkuΔβ*
D-truth better than D-falsehood
‘Truth is better than falsehood.’
6. *l-ša:y ʕa:li ħal-ay:a:m*
D-tea expensive these-days
‘Tea is expensive these days.’
7. *l-ʕanam ya:kil kil šay*
D-goats.COLL 3M.SG.eat every thing
‘Goats will eat anything.’

8. l-ʔe:s^s yighi:s^s w s-se:b yisu:b
 D-pearl-divers 3M.SG.dive and D-puller 3M.SG.pull
 'Pearl-divers would dive and the puller-men would pull.'

Proper nouns are divided into ones that take the definite prefix and ones that do not (the choice being lexically determined). Thus some place names take the definite article (e.g. *ʔal-ʕira:q* 'Iraq'; *ʔas-su:da:n* 'Sudan'), while others, including most borrowings, do not (e.g. *ʕu:man* 'Oman'; *ʔa:mrika:* 'America') ([11]:200).

Partitive quantifiers are always followed by a definite-marked noun phrase, whether semantically definite or indefinite (c.f. [13]:131-132):

9. a. ʔaʔlab r-ɾaya:yi:l
 most.of D-men
 'most (of the) men'
- b. l-kaθi:r min s-simaf
 D-whole.lot.of D-fish
 'a whole lot of fish'

Finally, demonstrative determiners can only modify definite-marked noun phrases independently of position:

10. a. ha:ði d-diri:ʕa
 this D-window
 'this window'
- b. d-diri:ʕa ha:ði
 D-window this
 'this window'

In contrast, in English, as the glosses indicate, abstract, mass, collective and generic nouns do not allow or require the definite determiner, while demonstrative determiners are always in complementary distribution with the definite determiner. Finally, proper names contain the definite determiner in only a few restricted contexts (usually geographical and landmark names). This difference in use predicts that the definite determiner in Emirati Arabic should appear significantly more frequently than the definite determiner in English. Although, corpus studies of Emirati (or Gulf in general) Arabic do not exist, we can get a rough percentage of definite marker use in Arabic from the child-targeted adult speech. We found that the percentage of definite determiner with respect to the full set of noun phrases that could potentially get a determiner is around 32%. On the other hand, in British English the definite determiner appears only in around 20% of the possible environments¹. This difference in frequency that is observed in the input, as we will see, may play a role in the acquisition of the definiteness system in the two languages.

3. THE DATA

3.1 The Database

The study is based on a set of monthly recordings of a group of six Emirati Arab children, for a period of six months (ages 1;6-4;4). The children come from a middle class socio-economic background with high school and university educated parents employed in government positions or staying at home. The language in their home environment is Emirati Arabic², a local variety of Gulf Arabic, while they are also exposed to a

¹ Longman Grammar of Spoken and Written English [17], states that there are about 20,000 definite determiners per million words in the conversation English corpus. This corresponds to around 100,000 singular nouns and 10,000 plural nouns. Thus, roughly 20% of English noun phrases in conversation are marked with the definite article.

² The term 'Emirati Arabic' is not attested in the relevant literature. It is used here as an administratively defined term to designate all minor varieties of Gulf Arabic spoken within the confines of the U.A.E. For a list of the distinctive grammatical characteristics of these varieties see [11], [13], and [19].

pidgin variety of Arabic, spoken by domestic helpers working in the house. The investigator visited the children’s house and recorded interactions between the children themselves as well as between the children and the investigator. The length of each recording was 30 minutes and took place at two-week intervals. After completion of the recording phase the project assistants transcribed the recorded session, using the Arial Unicode font that allows for IPA transcription of language specific sounds. Transcription was entered directly into a customized database platform using CLAN software which is freely available in the CHILDES language acquisition website. The transcription and coding format followed a simplified version of Codes for the Human Analysis of Transcripts (CHAT). All child and adult utterances were glossed in English.

Transcription was followed by the coding of the total set of noun phrases in the data. All noun phrases, including ones with null heads (cases of NP-ellipsis or substantivization) were coded. In addition, presence or absence of the definite marker was noted, with special attention to cases where the determiner is inserted or omitted erroneously. Finally, special codes were used for the presence of pronouns, demonstratives, proper names, and possession. A list of codes is provided in Table 1:

Table 1: List of Coding Conventions.

Code	Environment in Corpus
DP	Noun phrase
D	Presence of definite marker (ə)l-
0D	Absence of definite marker
W0D	Definite marker omitted erroneously
WD	Definite marker inserted erroneously
PRO	Pronoun
PN	Proper Name
DEM	Demonstrative
POSS	Possessive construction of any type

3.2 Results

We found that children from the early stages use definiteness markers productively, which seems to indicate that the D-system is intact. In the early stages (1;6-2;00) children almost exclusively use pronouns, demonstrative pronouns and bare nominal forms:

11. haaya babaah (Hind,1;10)
 this daddy
 ‘This is (my) daddy.’

In later stages of acquisition (2;7-4;4) the definite marker appears very productively and almost always in the right contexts:

12. haaya ?anuuta ?ətaleet-a t-taan (Al Reem, 2;9)
 this Hind bought-it D-grocery
 ‘It’s this Hind, I bought it (from) the grocery’.

However, the frequency of use of the definite marker seems to be lower than that in child-directed adult speech. The following graph compares the frequency of use in adult speech to the average frequency in child speech in three different stages of acquisition:

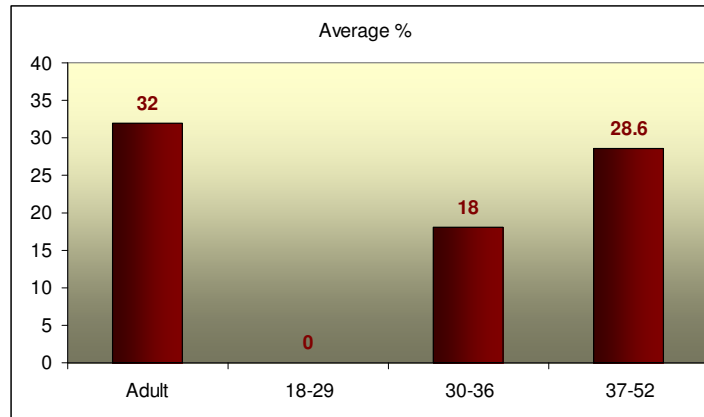


Fig. 1: Comparison of Adult and Child Definite Determiner Frequency

As can be seen in Figure 1, in adult speech there is an average of 32% of noun phrases with an overt definiteness marker out of all noun phrases that could potentially get a marker (i.e. excluding pronouns and determiner-less proper names). Children, on the other hand, produce no determiners in the early stages (18-29 months), while the frequency of determiners starts rising at around 30-36 months to reach almost adult levels during 37-52 months.

During these stages, children optionally omit the determiner in contexts in which it is obligatory in the target language (usually proper names or in front of generic nouns):

13. a. haaḍa laa ʕeemuul (Mohamed, 4;1)
 this not Al Ain Mall
 ‘This is not (from Al) Ain Mall’
- b. raaḥ-t merteen ʕiyaada (Abdulaziz, 3;6)
 went-3SG.FEM twice clinic
 ‘She went two times (to the) clinic.’

The percentage of omission (i.e. the error rate) is significant. The graph in Figure 2 illustrates determiner omission from obligatory contexts (out of all possible contexts):

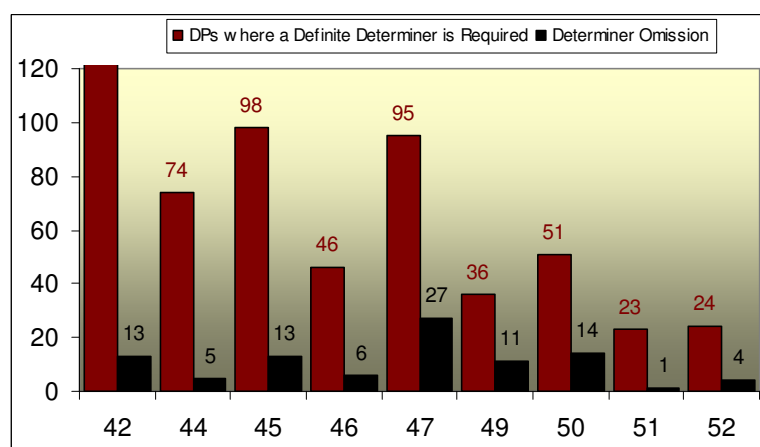


Fig. 2: Omission of Definiteness Markers

As can be seen, children often err when the determiner is required. The percentage of all omissions in the data is 14.92% of all potential environments.

Finally, in very few cases children use definite determiners in contexts which would be ungrammatical in adult language. These mostly include marking definiteness twice – once with the prefix and once with a suffixed possessive pronoun (as in 14 below):

14. b-aʕab luuħi **bə-l-liid-i** (TARGET: *b-iidi*) (Mohamed, 3;11)
 FUT-play.1SG myself by-D-hand-1SG.POSS
 ‘I will play by myself with my hand.’

However, these types of error are extremely rare in the data. There are only 35 attested cases in the database, out of a total of 1650 noun phrases (a percentage of 2.1%).

4. DISCUSSION

In most syntactic accounts pronouns involve a referential layer and therefore, they either merge in D [14] or move there from a lower syntactic projection ([15]). In addition, demonstratives either merge or move to the specifier of DP ([16] and references therein). Thus, both demonstratives and pronouns require the presence of a D head in the structure in order to be licensed. Given that the children from the very early stages use almost exclusively demonstrative and personal pronouns, it seems impossible to maintain the position that the D-layer is not present in child grammar. Therefore, the data provides strong evidence for theories of language acquisition that postulate that child grammar is equipped with the same set of functional projections as adult grammar from the initial stages.

The productive, and almost exclusive, use of demonstrative and personal pronouns in the early stages of language acquisition confirms the prediction in [9] that the development of referential expressions in child speech will follow the givenness hierarchy of cognitive statuses relevant to the form of referring expressions in natural language discourse. This is because pronouns and demonstrative pronouns are at least activated, i.e. the referent is represented in current short-term memory, a status that is expected to be prevalent in children of these ages. Use of definite determiners and indefiniteness markers signals uniquely identifiable and type identifiable referents and requires knowledge of addressee’s state of mind and knowledge about the world, burdening children with additional pragmatic constraints and thus, these markers are not used productively in the very early stages.

As we have seen, between the ages of 2;7-4;4 the definite marker appears very productively and almost always in the right contexts in child speech, but the percentage of occurrence is significantly lower than that in child directed adult language (see Figure 1). What could the reason for this mismatch between adult and child frequencies be? Remember (Section 2) that lack of the definiteness marker in Emirati Arabic marks an indefinite context (although this is a somewhat simplified picture, see [10] for discussion). This seems to indicate that Emirati children have a higher percentage of indefinite contexts in their speech than Emirati adults do. But this seems somewhat peculiar in that there is no obvious reason why it would be so and in fact contradicts findings in other languages (see for example [18] for Greek), which show clearly that indefinite noun phrases appear relatively late in the developmental stages.

A more possible answer is that Emirati children drop the determiner in environments where adults would have inserted it, resulting in a higher percentage of bare noun phrases in child speech. The immediate advantage of such an assumption is that the results in the acquisition of definiteness in Emirati Arabic would correspond to similar results in other languages that exhibit determiner drop and use of bare noun phrases. If this is on the right track, then a more careful re-examination of noun phrases in the child speech is required so that it can be established whether an adult would have used a marker that corresponds to a more restrictive cognitive status (i.e. a determiner or demonstrative). We leave this for further research (but see [17] for such work in Hebrew).

The important observation is that determiners are mostly omitted in environments in which they are syntactically required. As we have seen (examples 13.a-13.b), in a lot of cases the determiner is omitted when it forms part of a proper name, or before generic nouns. The following graph shows percentages of different environments where the definiteness marker is omitted:

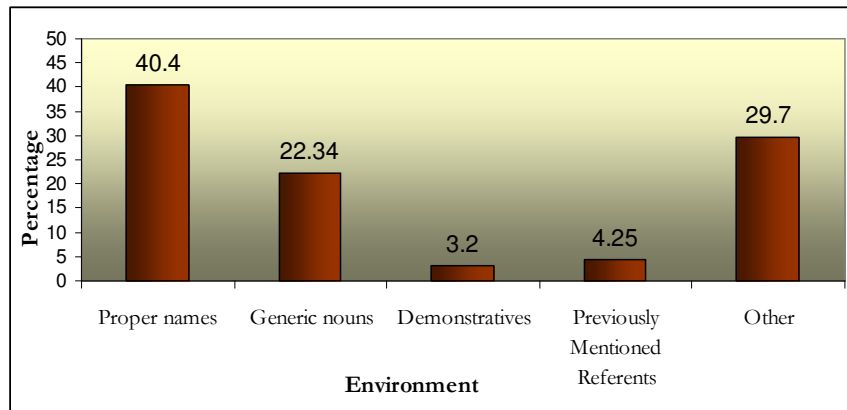


Fig. 3: Environments where the definiteness marker is omitted

As can be seen, the main environment (just below half of the cases) involves proper names which contain the definiteness marker, such as one of the children’s names (Al Reem), or place names in the region (Al Jimi and Al Ain malls). It seems then that children at these stages have not yet completely mastered the fact that the article is an integral part of the proper name and cannot be omitted.

A second environment where children predominately omit the determiner is that of generic nouns. We use the term ‘generic noun’ here to refer to entities in ‘ground information’, i.e. participants in the discourse that do not play a prominent role when first introduced. In all these cases, Emirati speakers mark the participants with a definite determiner. This mainly includes locations, such as ‘the school’, ‘the restaurant’, ‘the clinic’, but also human referents that may work in places like these, such as ‘the teacher’, ‘the doctor’, ‘the waiter’, and who play secondary roles in the discourse. On the other hand other inanimate referents with limited textual prominence are left unmarked when first introduced in the discourse (see [10]: 31-35 for discussion).

It seems then children at these stages have not yet mastered the complex interplay between syntactic environment and pragmatic/discourse knowledge, which is needed in order to place definiteness markers in the appropriate environments in their speech. However, this does not mean that children are not aware of the use of definiteness markers in the correct environments. Support for this comes from the extremely rare use of definiteness markers in environments where they are not permitted in the target language (e.g. example 14). As we have seen children make mistakes of this type extremely rarely, and as the following graph illustrates, most of these mistakes are again in the area of proper names, where children tend to add the definiteness marker to names that do not contain it:

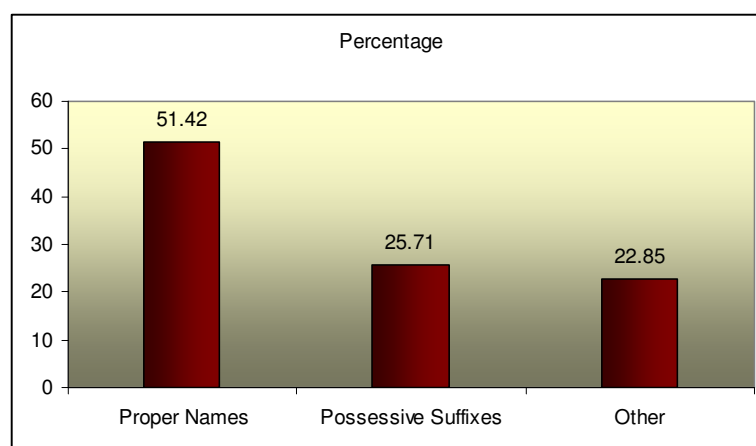


Fig. 4: Environments where the definiteness marker is inserted erroneously

The second most prominent environment is a case of double marking. As we have seen, in Section 2, the addition of a possessive suffix at the end of the nominal root in Emirati Arabic marks definiteness. In those cases a definiteness marker is not allowed. It seems then that children in the early stages have not fully mastered this rule and optionally allow double marking on their nouns.

In conclusion, it is interesting to note here that the section of Emirati Arabic child grammar that relates to the expression of definiteness does not vary from the target grammar in wild and unpredictable ways. All observations that have been discussed here relate to patterns of nominal expressions that are widely attested in other adult grammars. Take for example the cases of article omission in proper names, generic nouns (i.e. novel places introduced in the discourse) and noun phrases preceded by deictic demonstratives. In adult Emirati Arabic all these cases are marked with the definite determiner. In Emirati child language they resist definiteness marking for a short period until children master their distribution. In English they are either ungrammatical with a definite article or do not require it. Thus, Emirati Arabic child grammar behaves like English grammar for a short period of time, before adequate exposure to the input aids children to develop an adult-like grammar.

CONCLUSION

We have shown that there is a tension between syntactic and pragmatic factors in the acquisition of definiteness in Emirati Arabic. On the one hand, the correct use of the definite marker, as well as the use of proper names, pronouns and demonstratives supports the view that children have the full set of projections from the start. On the other hand, the rate of definite marker omission, as evidenced by the low frequency of the marker in child data and from the non-target-like use of the marker in certain environments like proper names and generic nouns indicates that the D-projection is underspecified and that children have not completely mastered the use of the marker.

One of the research objectives for subsequent work needs to be the careful examination of different noun phrases in child data with respect to the pragmatic contexts and cognitive statuses of the participants. In other words, while syntactically triggered errors have been established with satisfactory accuracy, it is not as yet sufficiently clear to what degree children have mastered the appropriate use of the definiteness marker in different pragmatic contexts.

ACKNOWLEDGEMENT

This work was financially supported by the Research Affairs at the UAE University under a contract no. 02-06-1-11/08. We are grateful to the children who participated in this study and their parents for granting their consent. Finally, we would like to thank Othaiba Tamimi for her work with transcribing, translating, and coding the data as well as her contribution to the write-up of the paper. Without her this project would not have been possible. Thanks also to Aamna Shemali and Fatima Dhaheri for working on the project.

REFERENCES

- [1] Radford, Andrew. 1990. *Syntactic Theory and the Acquisition of English Syntax: the nature of early child grammars of English*, Blackwell, Oxford
- [2] Radford, Andrew. 1996. Towards a Structure-Building Model of Acquisition, in *Generative perspectives on Language Acquisition*. Clahsen, Harald (ed). Amsterdam: Benjamins, p.43-89.
- [3] Rizzi, Luigi. 1994. Some notes on linguistic theory and language development: The case of root infinitives. *Language Acquisition* 3:371-393.
- [4] Wexler, Kenneth & David Poeppel. 1993. The full competence hypothesis of clause structure in early German. *Language*, 69:1, 1-33.
- [5] Hoekstra, Teun, Nina Hyams & Misha Becker. 1999. The Role of the Specifier and Finiteness in Early Grammar, in D. Adger, S. Pintzuk, B. Plunkett & G. Tsoulas (eds.) *Specifiers: Minimalist Approaches*. Oxford: Oxford University Press, pp. 251-270.
- [6] Gerken, L. & McIntosh, B. (1993). Interplay of function morphemes and prosody in early language. *Developmental Psychology*, 29, 448-457.
- [7] MacNamara, J. 1982. *Names for Things*. Cambridge, MA: The MIT Press.
- [8] Schaeffer, J. (2000). *The acquisition of direct object scrambling and clitic placement: Syntax and pragmatics*. Amsterdam/Philadelphia: John Benjamins.
- [9] Gundel, J.K., Ntelitheos, D. & Kowalsky, M. (2007). Children's use of referring expressions: Some implications for Theory of Mind. *ZAS Papers in Linguistics* 48, 1-21.
- [10] Brustad, Kirsten. 2000. *The Syntax of Spoken Arabic: A comprehensive Study of Moroccan, Egyptian, Syrian, and Kuwaiti Dialects*. Washington: Georgetown University Press.
- [11] Holes, Clive. 1990. *Gulf Arabic*. London, New York: Routledge.
- [12] Ryding, Karin C. 2005. *A Reference Grammar of Modern Standard Arabic*. Cambridge: Cambridge University Press.

- [13] Qafisheh, Hamdi A. 1977. *A Short Reference Grammar of Gulf Arabic*. Tuscon: The University of Arizona Press.
- [14] Abney, S. 1987. *The English Noun Phrase and its Sentential Aspect*. Doctoral Dissertation. MIT.
- [15] Zwarts, Joost. 1994. Pronouns and N-to-D Movement. In (M. Everaert, B. Schouten and W. Zonneveld (eds.) *OTS Yearbook 1993*, Utrecht: Utrecht University, 93-102.
- [16] Giusti, Giuliana. 1997. The Categorical Status of Determiners. In Liliane Haegeman, ed., *The New Comparative Syntax*. New York: Longman.
- [17] Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finegan. 1999. *Longman Grammar of Spoken and Written English*. Essex, UK: Longman.
- [18] Marinis, Theodoros. 2003. *The Acquisition of the DP in Modern Greek*. Amsterdam: John Benjamins.
- [19] Johnstone, T. M. 1967. *Eastern Arabian Dialect Studies*. London: Oxford University Press.