

Two Kinds of Definiteness in Coptic

Barbara Egedi¹

Abstract

The paper deals with a grammatical micro-variation that can be observed between Coptic dialects. All dialects have a rich system of determiners, but in certain varieties simple definiteness can be marked by two series of definite articles. According to the proposed hypothesis, in dialects that make use of a double system of determination, in Bohairic and in Fayyumic, the distribution of the articles corresponds to the strategy as to how the referent of a noun phrase is identified in the given discourse. The main claim is that weak articles grammaticalized to encode inherently unique and inherently relational referents, while strong articles are used in anaphoric contexts. This model will account for the asymmetry attested in plural forms as well as for the seemingly inconsistent variation of determiners in similar syntactic contexts.

1 Introduction

1.1 Aims

The aim of this paper is to describe in a comparative way how definiteness is marked in certain Coptic dialects. While the situation is well described in Sahidic, the distribution of definite determiners is less understood in Bohairic and Fayyumic, where simple definiteness is marked by two series of definite articles. The paper will overview this micro-variation, on the one hand, and aims to answer the question, on the other hand, what the functional difference is between the definite determiners in dialects that make use of a double system of determination. I will propose that the functional split can be explained by taking into consideration how the referent of a noun phrase can be identified in a given discourse (see section 2 on the notion of referential identification). Even though the hypothesis presented in this paper has not yet been widely tested, it aims to explain certain features in grammar that have not been adequately accounted for. Moreover, some observations that come from general linguistic literature, both synchronic and diachronic, may support the claims from a typological perspective as well.

1 Research Institute for Linguistics, Hungarian Academy of Sciences, Budapest (name[at]email.com).

Work on this paper was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences, and by the National Research, Development and Innovation Office (OTKA/NKFI 112828) which is gratefully acknowledged by the author.

After a short introduction about the definite article in Egyptian and about the Coptic dialects, the notions of definiteness and referential marking will be addressed in Section 2, also providing two semantic models that are able to handle the various uses of definite articles. Section 3 gives a survey of the article system in Sahidic, while Section 4 presents the Bohairic system and offers a proposal as to how the two kinds of definite articles can be explained in terms of the different strategies of referential identification. Some preliminary studies on early Fayyumic will be presented in Section 5, and finally, Section 6 concludes the paper with some typological considerations and some thoughts on language change.

1.2 The definite article in Egyptian

Earlier Egyptian had no definite article. The systematic and consistent use of the article in all written registers can be dated to the New Kingdom,² admitting that the definite article must have emerged (or more precisely, grammaticalized) as early as in colloquial Middle Egyptian. The Egyptian definite article – just like in many other languages – developed out of a series of demonstratives, the so called *p3*-series. Its emergence was one of the most significant changes in the history of the language, which also motivated to make a typologically distinction between the Earlier and the Later phase of Egyptian (as suggested by Loprieno 1995). The emergence of the definite article had a more general consequence that concerns the complete reorganization of nominal constructions, but this latter topic falls out of the scope of the present paper.

What factors conditioned the change in definiteness marking seems to be an understudied question in the literature. Not much attention has been given either to the early history of the definite article, or to the changes in its functions in the subsequent language stages. As to my knowledge, a detailed description of the process was only provided by Kroeber (1970) and Loprieno (1980). Recently, Kupreyev (2014) and Zöllner-Engelhardt (2016) approached this linguistic problem in a more detailed manner.

It must also be noted that while the more or less standardized written traditions of the language stages before Coptic concealed the possible geographical divergences, many differences can be attested in the article systems of the Coptic dialects, as will be discussed in the following sections.

1.3 The Coptic dialects

The Coptic dialects have been claimed to differ mainly in pronunciation and spelling, in addition to minor diversities in vocabulary. The majority of grammatical features that might appear to be divergent between dialects manifests as formal varieties conditioned by general phonological rules and morphophonological correspondences (cf. Funk 1991).³ Recently, more efforts have been done to demonstrate cases of micro-variation in syntax as well. It must be remembered, however, that there are no coherent syntactic descriptions

2 For definite and possessive articles in Late Egyptian, consult: Černý-Groll (1978: §3.2 and §3.5); Erman (1933: §§171–182); Junge (1996: §2.1.1–3); Loprieno (1995: 69); Neveu (1996: §2.1).

3 About the Coptic dialects in general: *inter alia* Funk (1988); Kahle (1954: 193–278); Kasser (1991); Vergote (1973: 53–59); Worrell (1934: 63–82); and Egedi (2012: 17–28).


for all the dialects attested between the 4th and the 6th centuries, except for Sahidic and Bohairic, whose significance and use in the written records goes beyond that period.⁴ Minor literary dialects, unfortunately, disappeared in later sources, thus a comparative study will always be restricted in time and will also be subject to chance as to how many and what type of manuscripts have been preserved. Due to the gaps in syntactic descriptions, the inter-relationship of some dialects is still a matter of debate; therefore, addressing even minor grammatical features in a systematic way may contribute to the understanding of how closely certain varieties (and subvarieties) are related.

As for the use of the definite article, many of the minor dialects seem to behave like Sahidic, while other varieties (Bohairic and Fayyumic) definitely diverge since they make use of two sets of definite articles. This latter phenomenon will be focused on in the second half of the paper with the aim of exploring how these systems of determination worked as compared to Sahidic.

2 About definiteness and reference

The semantic and pragmatic notion of definiteness can be considered universal, its grammatical realization, however, is a language specific property.⁵ The definite article is a possible grammatical device to encode definiteness, undoubtedly a typical one, but it is far from being exclusive. Many languages have no article at all, since definiteness can be marked through case distinction, through aspect, through the position of the noun phrase, or the combination of more than one of the possible devices (van Gelderen 2011: 146). Furthermore, the extended use of possessive suffixes can easily fulfill the functions of determination. This strategy has been observed in several Uralic languages, as well as in Turkish (see *inter alia* Schroeder 2016: 585–598). At the same time, in languages that do have a definite article, the articles may behave quite differently from one language to another. There are also languages that may express definiteness by more (types of) determiners through distinguishing more than one set of definite articles.

The question that should be addressed first of all is what definiteness means in grammar. The basic function of the definite article (or any other grammatical strategy that encodes definiteness) is to identify the referent of the noun phrase: the speaker signals that the hearer is able to assign a referent for a certain noun phrase, either because it is accessible in the discourse, or because it is familiar to the hearer based on his/her general knowledge of the world.⁶ The existence and uniqueness of a referent referred to by a definite description must hold within the universe of discourse, which can be characterized by specific pragmatic parameters. This pragmatic set can even be very small (the immediately observable or just mentioned objects), or else, it can also be a considerably large set that

 Texts often usually provide an introductory chapter with notes on the grammatical peculiarities of the manuscript they publish. These observations, however, principally fall into the above mentioned category of phonological and morphophonological variation.

5 For a general overview on definiteness, see Lyons (1999, esp. Chs. 1. and 7) and Abbott (2004).

6 The idea of considering definiteness as referential identification principally follows Lyons (1999).

includes all entities whose existence is universally accepted. The point is that the speaker and the hearer must share the relevant set in their discourse situation (cf. Hawkins 1991).

The table below overviews the different ways how discourse referents can be identified in a conversation. The table is the result of a combined application of the models proposed by Himmelmann (1997, 1998, 2001) and Lyons (1999) who themselves relied on Hawkins (1978).⁷ The strategies for referential identification are illustrated through mini-discourses given in English, in the right column of the table:

Table 1 | Strategies for referential identification

| | |
|---|---|
| Anaphoric use | “There’s <i>a new café</i> in our street. We have nothing in the fridge. Let’s go and see what we can get in <i>the café</i> .” |
| Associative-anaphoric use | “There’s <i>a new café</i> in our street. <i>The cheesecake</i> is just perfect!” / “ <i>The owner</i> has just come back from Japan.” |
| Situational use = situational uniqueness | “What does <i>the cheesecake</i> cost?” (here, in the café where we are at the moment) |
| Larger situational use = inherent uniqueness | “ <i>The sun</i> is shining brightly.” “I am never bored with <i>the Italian cuisine</i> .” |

Anaphoric strategy is used when the object has been mentioned previously, like *the café* in the first row of Table 1. In associative-anaphoric use, the object itself was not concretely mentioned in the preceding textual context, its presence or existence is only assumed by association, as is the case with *the cheesecake* and *the owner* in the second row of the table. The situational use of definiteness means that the referent of the noun phrase is considered unique in the speech situation. In the above cited sentence “What does the cheesecake cost?”, the referent of *the cheesecake* can only be interpreted as unique if it is clear that we are speaking about the cheesecake that is sold in the café where we stay at the moment. Finally, inherently unique nouns are considered unique according to our knowledge of the world, so their referents can be identified independently of the speech situation.

What is common in inherent uniqueness, situational uniqueness and associative-anaphoric use is that they perform an *extra-linguistic* identification of reference. Whereas situational uniqueness is evidently not independent of the direct situation, inherent uniqueness is completely discourse-independent. The associative-anaphoric use is a more complex phenomenon: the referent of the noun phrase is identified anaphorically, but not in a direct way. The hearer (the reader), within a discourse situation, always activates a general, extra-linguistic knowledge in order to set a whole universe of possible referents which may be associated with the referents that were already introduced explicitly. Accordingly, the associative-anaphoric use is related to the larger situational use in another feature, since a *general knowledge of the world* is definitely needed for both, due to the fact that the referents are not present either in the previous discourse, or in the immediate speech situation. This feature has an important grammatical consequence as well: it has been universally observed that demonstratives can never be used in these two contexts.

7 For another approach with quite similar distinctions, consider also the three types of mental structures for grounding referents in Givón (2001: 459–465).

In the second part of the paper, the analysis will also rely on the semantic model of Sebastian Löbner (1985, 2011), who studied the interaction of determination with lexical meaning. Löbner distinguishes four basic conceptual lexical types of nouns, his proposal is summarized in Table 2 below. Sortal nouns are prototypical nouns, which are compatible with all modes of determination. Individual nouns include proper names, personal pronouns, nouns for unique institutions, abstract nouns, etc., which can be assigned a unique referent in every appropriate context. The referents of relational nouns are characterized in terms of their particular relation to some other object, typically a possessor expression (e.g. kinship terms, terms for non-unique parts, deverbal nouns, etc.). Finally, functional nouns are relational nouns that are, at the same time, unique (typically kinship terms, body parts, abstract terms, dimensions, such as *mother*, *author*, *head*, *age*, *price*, etc.).⁸ The lexical types can be characterized by two binary features: inherently relational types are marked as [+R], while inherently unique types are marked as [+U].

Table 2 | Basic conceptual lexical types of nouns (Löbner 2011)

| | –U | +U |
|----|--|---|
| –R | sortal nouns <e,t> <i>stone, book, adjective, water</i> | individual nouns <e> <i>moon, weather, date, Maria</i> |
| | | |
| +R | relational nouns <e,<e,t> <i>sister, leg, part, attribute</i> | functional nouns <e,e> <i>father, head, age, subject</i> |
| | | |

The default use of inherently unique [+U] nouns is singular definite. Löbner (2011) claims that inherently unique concepts are semantically unique, while non-unique [–U] concepts can only be pragmatically unique. Therefore he distinguishes congruent vs. incongruent definiteness. Extra marking of semantic/congruent definiteness is redundant in Löbner’s model. This claim predicts that an asymmetry may easily develop in the grammatical encoding of semantic and pragmatic uniqueness. As will be shown, this prediction nicely corresponds to what we find when observing the distribution of Coptic articles in dialects with a double article system.

3 Definite articles in Sahidic

Examining the Sahidic system of definite articles might be an appropriate point of departure before turning to comparative considerations. The Sahidic articles present a couple of context-dependent allomorphs, as shown in examples (1) and (2) below:

- (1) a. **ⲡⲣⲱⲙⲉ** b. **ⲧⲥⲱⲩⲉ** c. **ⲛⲣⲱⲙⲉ/ⲛⲥⲱⲩⲉ**
 p-rôme t-sôše n-rôme/n-sôše
 DEF.SG.M-man DEF.SG.F-field DEF.PL-man/DEF.PL-field
 ‘the man’ ‘the field’ ‘the men/the fields’

⁸ Note that nouns are usually polysemous, so they can belong to more than one types, or undergo a type-shift in a given utterance.

- (2) a. **πεζροογ**
pe-hroou
DEF.SG.M-voice
‘the voice’
- b. **τεζσιμε**
te-shime
DEF.SG.F-woman
‘the woman’
- c. **νεσνιγ**
ne-snêu
DEF.PL-brother
‘the brothers/siblings’

The articles are always proclitic and display the grammatical features of gender and number just as other determiners. This means that they have three allomorphs depending on the noun they are associated with – even though the noun itself does not present these morphological categories, except for irregular cases (e.g. **σνιγ** in (2c)). Accordingly, the articles have a masculine and a feminine form in singular, and a gender-independent plural form. The longer forms **πε-**, **τε-**, **νε-** in (2) are conditioned by phonological constraints. Long articles are regularly attested before consonant clusters (more precisely before a complex syllable onset). Longer forms also appear with a few nouns denoting time expressions, such as **πε-υοεισ** pe-uoeiš ‘the time’, **πε-ζροογ** pe-hroou ‘the day’, **τε-ρομπε** te-rompe ‘the year’, **τε-γνογ** te-unou ‘the hour’, **τε-γση** te-ušê ‘the night’.⁹

Definite determiners, including articles are collected in Table 3 below. They form a natural class in the sense that they mutually exclude each other and are always interpreted as definite. The determiners in the first two rows are usually considered simple articles without encoding deixis or possession.

Table 3 | Definite determiners in Sahidic

| | SG.M | SG.F | PL |
|-------------------------|-------------------------|-------------------------|-------------------------------|
| Definite article | π-/πε- p-/pe- | τ-/τε- t-/te- | ñ-/ñ-/νε- n-/n-/ne- |
| π -determination | π - pi- | τ - ti- | ñ - ni- |
| Demonstrative article | πει- pei- | τει- tei- | νει- nei- |
| Possessive article | πεq- pef- | τεq- tef- | νεq- nef- |

In what follows, the so called **π**-determination will shortly be discussed. The function of this series is not completely understood as sometimes it seems to be closer to demonstratives. As a matter of fact, the **π**-series in Sahidic has often been described as the reduced form of the demonstrative article.¹⁰ It was Hans Jacob Polotsky (1957: 229–230), who first listed exhaustively the four contexts in which the **π**-series regularly appears:

9 For a full *Formenlehre* of Sahidic definite articles, consult the following grammars: Lambdin (1983: §§1.3, 17.2); Layton (2000: §52); Steindorff (1951: §136–139); Stern (1880: §§227–230); Till (1961: §§87–91, 94–99); Vergote (1983: §§121–122, 124–125).

10 For the determiner **π**- and its corresponding pronominal form **πñ**-, see Lambdin (1983: 30.8, *remote demonstratives*); Layton (2000: §58, *affective demonstrative*); Steindorff (1951: §89 and §136); Stern (1880: §§227); Vergote (1983: §§127–128).

- i.) It appears as a fixed component in temporal and spatial expressions, in which its function can clearly be derived from an earlier demonstrative meaning.
- ii.) It appears in comparative expressions of the form **ⲡⲉⲛⲉ** **ⲡⲉⲛⲓ**- *nt^{he} n-ni* ‘like’, but always in plural form.
- iii.) It appears as an anaphoric identity marker¹¹ in the reinforced expression **ⲡⲓ**... **ⲡⲓ-ⲟⲩⲱⲧ** *pi- ... n-ouôt* ‘the same...’.
- iv.) Finally, the **ⲡⲓ**-series may have an affective/emotive use in attributive constructions, when something is described as admirable or horrible.¹²

The use of **ⲡⲓ**-determination does not seem to be frequent in Sahidic. For example, apart from its appearance in set phrases (e.g. **ⲉⲡⲓⲕⲣⲟ** **ⲛ**- *e-pi-kro n-* ‘across/beyond’), **ⲡⲓ**- is only attested twice in the 5th century Gospel of John (P. Palau Ribes Inv 183, used as a test-corpus in my earlier studies):

(3) John 5:44, Sahidic (S)

| | | | |
|----------------|--------------------------|----------------|----------------|
| ⲡⲉⲟⲩⲱ | ⲡⲉⲃⲟⲗ ⲉⲓⲧⲙ | ⲡⲓⲟⲩⲁ | ⲡⲓⲟⲩⲱⲧ |
| p- <i>eooy</i> | p- <i>ebol hitm</i> | pi- <i>oua</i> | n- <i>ouôt</i> |
| DEF.SG.M-glory | DEF.SG.M-out from | DEF.SG.M-one | ADJZ-single |

‘(How will ye be able to believe, taking glory from one another, and) the glory which is from the/this only one (ye seek not for).’¹³

(4) John 9:11, Sahidic (S)

| | | | | |
|-------------------|-------------------------|---------------------|---------------|------------------------|
| ⲁⲩⲟⲩⲱⲃⲉ | ⲭⲉ ⲡⲓⲣⲱⲙⲉ | ⲉⲧⲟⲩⲙⲟⲩⲧⲉ | ⲉⲣⲟⲩ | ⲭⲉ ⲓⲥ |
| a-f- <i>ouôšb</i> | če pi- <i>rôme</i> | et-ou- <i>moute</i> | ero- <i>f</i> | če i- <i>êsou>s</i> |
| PF-3SG.M-answer | that DEF.SG.M-man | REL-3PL-call | to-3SG.M | that Jesus |

‘he answered that the/this man who is called Jesus (is he who made clay and put it on my eyes...)’

The meaning of the sentence in (3) suggests a literal translation of **ⲟⲩⲱⲧ** ‘single’ and, consequently, the so called emotive use of **ⲡⲓ**- might be attested here. The quote in (4) forms part of a *cleft-sentence*. The speaker here does not simply answers a question, but also wants to focalize the agent of the event. This might be the reason for his using an affective determiner.¹⁴

As will be shown in the following section, the Bohairic determiners that correspond in form to the Sahidic emotive article, have quite different functions.

11 The term is mine. Cf. also Layton’s (2000: §58) definition of this use as “insisting upon identification”.

12 The use of an emotional article is not unparalleled: it has been reported, for instance, in certain Polynesian languages to convey sympathy or belittlement, its main function being to add information about the speakers’ attitude. Cf. Himmelmann (2001: 836), with references to the relevant literature.

13 The Coptic examples are cited after Quecke (1984). The translations are drawn from Horner’s edition of the Gospel with the relevant context.

14 Note that the Lycopolitan London Gospel (Thompson 1924) has a simple definite article at the locus cited in (3), while Horner’s Sahidic version and the version in dialect W have a demonstrative **ⲡⲉⲓ**- at the place cited in (4).

4 Definite articles in Bohairic

Bohairic demonstrative and possessive articles do not present any peculiarity as compared with Sahidic. At the same time, definite articles appear to constitute a real double system, with two series of definite articles. The two series are traditionally called ‘weak’ and ‘strong’ articles (see Table 4.). The longer forms (**πϵ-**, **τε-**, **νε-**) attested in Sahidic are absent here, but aspirated allomorphs of the ‘weak’ series appear before sonorant consonants with a relative consistency.¹⁵

Table 4 | Definite determiners in Bohairic

| | SG.M | SG.F | PL |
|-----------------------|-------------------------------------|-------------------------------------|---------------------|
| ‘Weak’ article | π-/ϕ- p-/p ^h - | τ-/θ- t-/t ^h - | nen- nen- |
| ‘Strong’ article | πi- pi- | ti- ti- | ni- ni- |
| Demonstrative article | πai- pai- | tai- tai- | nai- nai- |
| Possessive article | πεq- pef- | τεq- tef- | νεq- nef- |

(5) Early Bohairic (B4)

- | | | |
|--|--|--|
| a. πιομ p-iom DEF:SG.M-sea ‘the sea’ | b. τϕϵ t-p ^h e DEF:SG.F-sky ‘the sky’ | c. nenϥηρι ναβρααμ nen-šêri n-abraam DEF:PL-son POSS-Abraham ‘the sons of Abraham’ |
|--|--|--|

(6) Early Bohairic (B4)

- | | | |
|--|---|--|
| a. πiρωμι pi-rômi DEF:SG.M-man ‘the man’ | b. tiçimi ti-shimi DEF:SG.F-woman ‘the woman’ | c. niρωμι ni-rômi DEF:PL-man ‘the men’ |
|--|---|--|

Observe that the form **n-** seems to be entirely absent in Bohairic (but see Polotsky (1968) for a revision). The form of the plural ‘weak’ article is rather **nen-**, but its use is quite restricted since it only appears on possessed nouns, as illustrated in (5c). In all other cases, plural noun phrases are determined by **ni-**. This asymmetry between the singular and the plural forms, of course, needs to be accounted for.

According to Leo Depuydt (1985: 51) the two sets “cannot be studied regardless of their syntactical links with the two ‘genitive’ particles **n-** and **nte-**” because it is the possessive construction that conditions the determination of the first noun in the pattern. This claim,

¹⁵ The Bohairic examples are early Bohairic, cited from P. Bodmer III (ed. by Kasser 1958). Example in (5c) can be found at John 8:39. On classical Bohairic articles, see: Stern (1880: §§226–227); Mallon (1907: §§41–44); Polotsky (1968: 243); Depuydt (1985); Shisha-Halevy (1994; 2007: 430–447).

however, is an overgeneralization. In reality, only one of the possessive constructions is restricted in use, namely, the construction connected by **ⲛ**:- it has selectional criteria with respect to the lexical properties of the head noun which must be an inherently relational noun. As for the articles, the ‘weak’ series can appear in both possessive constructions, but in the construction mediated by **ⲛ**- only ‘weak’ articles can determine the possessed noun (Egedi 2012). At the same time, both ‘weak’ and ‘strong’ articles in singular appear independently, outside possessive constructions, and these uses have to be accounted for as well.

Of course, the issue has already been discussed by many authors, who aimed to explain the functional difference of the two sets of Bohairic articles. The history of research is quite long (see below). Even though the claims are not necessarily conflicting in the literature, the terminology is not uniform. Therefore a variety of definitions will be cited in Table 5 to illustrate the different opinions. After evaluating this table, a new proposal will be presented which aims to cover all the phenomena that can be observed in this dialect and which is based on distinguishing semantic vs. pragmatic definiteness. To support the hypothesis, a whole paragraph will be cited from the *Gospel of John* with an analysis that explains the distribution of the determiners.¹⁶

Table 5 | Earlier definitions of Bohairic articles

| | ‘Weak’ articles | ‘Strong’ articles |
|--|---|--|
| Stern (1880: §227) | “(…) hat der schwächere artikel gewöhnlich eine allgemein bestimmende bedeutung, und findet daher vor generischen und abstracten begriffen seine anwendung” | “(…) hat der starke artikel eine vereinzelt und unterscheidend bestimmende bedeutung” |
| Mallon (1907: §42) | “détermine d’une manière moins précise; il se place devant les noms génériques ou abstraits et devant les noms d’êtres uniques” | “détermine avec plus de précision, il indique un individu en particulier” |
| Polotsky (1968: 243) | ‘generically’ | ‘individual’ |
| Depuydt (1985: 59) | “– the indication of unique beings, – the generic use, – the use ‘par excellence’ (e.g. the river), which all derive from the basic notion of indicating one element of a genus as the representative of the entire genus” | <i>agrees on previous definitions</i> |
| Shisha-Halevy (1994: 233–234) and (2007: 389, 392) | “the genus or class naming determination” “the determinator is deictically inert, non-phoric, properizing” “non-cohesive, pure actualization designative or naming article” | “deictic, cohesive specifying article” “characterizes the noun as familiar and of high specificity” |

16 The textual context is extremely important when examining strategies of determination; in this case isolated sentences are not informative enough.

Earlier works cited here claim that weak articles have a generic function and determine less precisely than strong articles do. Strong articles, on the other hand, indicate individual or particular reference. Mallon's definition already refers to the concept of *uniqueness*, which is a central notion in this paper. From Shisha-Halevy's complex explanation, it is worth picking out the *non-phoric* nature of weak articles, which is also in accordance with what will be proposed below.

Based on the models presented in section 2, I propose to explain the distribution of the two series of article in the following way. Weak articles can be found with inherently unique nouns, both concrete and abstract nouns, whose referents can be identified through a general knowledge of the world, or else, through the knowledge of the actual situation. This would correspond to the strategies called situational use and larger situational use in section 2. Weak articles also appear with singular generic noun phrases, because generics refer to kinds, and as such, they behave like labels, and are usually associated with an existential presupposition. Strong articles are anaphoric instead, that is to say, they are used in contexts where the referent of a noun phrase can only be identified in the given discourse. Accordingly, strong articles encode pragmatic definiteness in anaphoric and associative-anaphoric contexts.


This proposal also accounts for the asymmetrical picture that has been observed in the case of plural noun phrases. Weak articles appear in possessive constructions as determiners on the possessed noun. Moreover, the plural weak article **nen-** can only appear in possessive constructions. This might allow  to assume that weak articles grammaticalized not only to encode inherently unique concepts, but to encode inherently relational concepts as well. Table 6 repeats the original table from Löbner (2011), cited as Table 2 above, with the gray cells marking the functional domain where weak articles appear in Bohairic.

Table 6 | Basic conceptual lexical types and grammaticalized weak articles in Bohairic

| | -U | +U |
|----|--|---|
| -R | Sortal nouns <e,t> <i>stone, book, adjective, water</i> | Individual nouns <e> <i>moon, weather, date, Maria</i> |
| +R | Relational nouns <e,<e,t> <i>sister, leg, part, attribute</i> | Functional nouns <e,e> <i>father, head, age, subject</i> |

From this proposal it naturally follows that weak articles either appear in singular with inherently unique nouns in situational or larger situational use, or they appear in both singular and plural in possessive constructions. The reason for the fact that the plural weak article **nen** never appears outside possessives is straightforward: inherently relational nouns may appear both in singular and plural, but an inherently unique concept will necessarily figure in a *singular* noun phrase.

It might be useful to stop for a moment and to raise the question how we know that anaphoric, strong articles are indeed articles, rather than demonstratives. Marking discourse-anaphoric reference is one of the main functions of demonstratives as well. Moreover, as was indicated above, **m**-determination in Sahidic still preserves some

features of the original demonstratives from which it developed. However, Bohairic strong articles are different from the Sahidic **ⲙ**-series. Bohairic strong articles are also used in associative-anaphoric contexts, in which no demonstrative can normally appear (cf. Section 2 for this claim). This use is illustrated in (7) below:

(7) John 11:17, Early Bohairic (B4)

| | | | | | |
|---|-----------|-----------------|------------|-------------------|----------------------|
| ⲁⲩⲓ | ⲁⲉ | ⲛⲭⲉ | ⲓⲏⲥ | ⲁⲩⲭⲉⲙⲓ | ⲉⲁⲩⲟⲩⲱ |
| a-f-i | de | nče | iê<sous>s | a-f-čem-f | e-a-f-ouô |
| PF-3SG-come | SP | PTCL | Jesus | PF-3SG-find-3SG | SBRD.PF-3SG-complete |
| ⲉⲩⲉⲣⲁ | | ⲛⲛⲉⲥⲟⲩⲱ | | ⲫⲉⲛ-ⲙⲓ-ⲙⲁⲩ | |
| e-f-er-4 | | n-n-ehoou | | xen-pi-mhau | |
| SBRD-3SG-do-4 | | PART-DEF.PL-day | | in-DEF.SG.M-tomb | |
| ‘Jesus came, and he found that he (Lazarus) had already been in the tomb four days.’ | | | | | |

In this sentence, “the tomb” is mentioned for the first time, but the noun phrase is determined by a strong article. The directly preceding context reports about the death of Lazarus and about the fact that Jesus wants to see the sight it happened. The referent of “the tomb” is not accessible either in the situation or in the previous discourse. It is definite, because its existence is taken for granted through association. This is a typical associative-anaphoric context, from which demonstratives are excluded by nature. Thus the determiner **ⲙ** in examples such as (7), must be an article.

Nevertheless, the distribution of the articles in the Coptic manuscripts does not appear as regular as it might be expected on the basis of the picture outlined above. A great oscillation, an apparently nonsystematic variation can be observed in the data. The same noun or noun phrase sometimes appears with a weak article, sometimes with a strong article. This overlap, however, can be derived from the twofold approach proposed here. The use of weak articles has been defined by semantic criteria, while the use of strong articles is purely defined by pragmatic criteria. Variation follows from the fact that such a system is very permissive with respect to the pragmatic definiteness. A concept which is unique or relational in a semantic sense, can always become anaphoric in a given discourse through the mere fact that it is mentioned repeatedly. This means that pragmatic factors can freely overwrite the basic conceptual types, and, at the end, it is the chosen strategy of referential identification that will select the appropriate article.

Examining sentences in isolation is hardly ever informative enough, as far as referential properties are concerned. For this reason, a longer text sample has been chosen from the early Bohairic Gospel of John (P. Bodmer III. edited by Kasser 1958), where most types of referential identification can be studied. For sake of simplicity, glosses are not given for the whole excerpt, but the relevant noun phrases are high-lightened both in the original Coptic text and in the English translation. An analysis of the determined noun phrases follows the citation.

(8) John 6:16–21, Early Bohairic (B4)

ΕΤΑΡΟΥΣΙ ΔΕ ΨΩΠΙ ΑΝΕΓΜΑΘΗΤΗΣ ΕΙ ΕΖΡΗΙ ΕΦΙΟΜ · ΑΥΑΛΗ ΕΥΧΟΙ ΑΥΨΕ
 ΕΠΑΤ ΜΠΙΟΜ ΕΚΑΦΑΡΝΑΟΥΜ ΝΕΑΤΧΕΜΤΣ ΟΥΩ ΕΣΨΩΠΙΠΕ ΝΕΜΠΑΤΕΙΗ̅̅̅ Ι
 ΖΑΡΨΟΥΠΕ · ΝΕΑΠΙΟΜ ΔΕ ΤΨΟΥΝ ΕΨΩΠΙΠΕ ΝΤΕΝΟΥΧΙΝΝΙΚΙ ΝΤΕΥΝΙΨΤ
 ΝΘΟΥ · ΕΤΑΥ<ΟΥ>ΕΙ ΕΒΟΛ ΝΚΕΝΣΤΑCION ΙΕ λ̅ · ΑΥΝΑΥ ΕΙΗ̅̅̅ ΕΦΜΩΠΙ
 ΖΙΧΕΝΠΙΟΜ ΕΑΦΨΩΝΤ ΕΠΧΟΙ · ΑΥΕΡΖΟΤ · ΝΘΟΥ ΔΕ ΠΕΧΑΔ ΝΨΟΥ
 ΧΕΑΝΟΚΠΕ ΜΠΕΡΕΡΖΟΤ · ΝΑΥ<ΟΥ>Ψ ΔΕΠΕ ΕΨΟΠΙ ΕΡΨΟΥ ΕΠΙΧΟΙ
 ΣΑΤΟΤΥ ΑΠΙΧΟΙ ΜΟΝΙ ΕΠΧΡΟ ΕΠΜΑ [ΕΝΑΥ]ΝΑΖΩΛ ΕΡΟΔ

“¹⁶An evening having come, his disciples came down to **the sea**; ¹⁷and having entered into **a ship**, they were going across **the sea** to Kapharnaum. And **the dusk** had now come, and Jesus had not yet come to them. ¹⁸And **the sea** was heaving by the blow of a great wind. ¹⁹Having then been distant about twenty-five stadia or thirty, they saw Jesus walking upon **the sea**, approaching **the ship**, and they feared. ²⁰But he said to them: ‘It is me, do not be afraid.’ ²¹They were wishing then to *get* him into **the ship** with them, and immediately **the ship** landed at **the shore** on **the land** to which they were to go.”

The following list contains all the noun phrases of the text that are determined by a definite or an indefinite article. Each Coptic form is accompanied by its transcription, its translation and by the strategy with which the referent was identified. The strategy, of course, can only be defined in the knowledge of the textual context.

| | | | | |
|--------|-----------------------|----------------|-------------------------------------|----|
| ΦΙΟΜ | p ^h -iom | ‘the sea’ | first mention, but unique reference | ✓ |
| ΟΥΧΟΙ | ou-čoi | ‘a ship’ | first mention, new referent | ✓ |
| ΠΙΟΜ | pi-iom | ‘the sea’ | unique reference, anaphoric | ✓✓ |
| ΤΧΕΜΤΣ | t-k ^h emts | ‘the darkness’ | abstract noun, unique reference | ✓ |
| ΠΙΟΜ | pi-iom | ‘the sea’ | unique reference, anaphoric | ✓✓ |
| ΠΙΟΜ | pi-iom | ‘the sea’ | unique reference, anaphoric | ✓✓ |
| ΠΧΟΙ | p-čoi | ‘the ship’ | <i>expected</i> : anaphoric | ? |
| ΠΙΧΟΙ | pi-čoi | ‘the ship’ | anaphoric | ✓✓ |
| ΠΙΧΟΙ | pi-čoi | ‘the ship’ | anaphoric | ✓✓ |
| ΠΧΡΟ | p-k ^h ro | ‘the shore’ | relational | ✓ |
| ΠΜΑ | pi-ma | ‘the place’ | cataphoric | ✓ |

The noun phrases that are mentioned more than once in the paragraph deserve a special attention here. The first mention of “the sea” has a weak article, as its referent is unique in the situation. In its second, third and fourth mention, “the sea” displays a strong article because of the anaphoric use of the same noun phrase (anaphoric use is marked by double check-marks in the right column of the list). The referent of “the ship” is neither inherently, nor situationally unique, so it has an indefinite article when mentioned for the first time. However, after having been introduced into the discourse, it is determined by the anaphoric, strong article. The only exception to this tendency can be found in John 6:19, marked by a question-mark in the list. The referent of “the ship” is expected to be identified anaphorically, but it is the weak article that determines the phrase, in spite of the

fact that neither “the ship” is unique in this context, nor it is a relational concept, like e.g. **ⲡⲓⲭⲣⲟ** “the shore”, later in the text.¹⁷

5 Definite articles in early Fayyumic

Fayyumic is not among the most studied dialects of Coptic, in spite of the relatively long period that its sources cover. A sketchy grammatical description has been provided by Till (1930), where observations are mostly based on classical Fayyumic texts.

In a previous comparative study of possessive constructions in Coptic dialects (Egedi *forthcoming*), I argued that the early Fayyumic grammar of possessives is quite similar to that of early Bohairic. Evidently, this observation pushed me to examine whether the model proposed for Bohairic in the last section can be applied to the system of determination in this dialect. Till’s *Chrestomathie* (1930: 3) is laconic, simply listing the possible forms of articles with no interpretation. The nature of the only reference I found about the use of Fayyumic articles will clearly show how insufficiently this dialect has been described: in a footnote of his monograph on Bohairic syntax, Ariel Shisha-Halevy (2007: 387 n.28) quotes a letter from 2000, written by Wolf-Peter Funk, in which the latter scholar remarks that early Fayyumic (F4) is close to Bohairic as far as the plural article usage is concerned, while in the singular the situation is similar to that of Mesokemic.

The neglected status of the dialect has good reasons, of course. Early Fayyumic texts are few, and they are all very fragmentary. In a preliminary study presented here, I examined a single manuscript of the British Museum (BM Or. 5707, ed. by Crum and Kenyon 1900) that contains a short section from the *Gospel of John* (3:54:49). In this text 122 simple, definite determiners could be identified. According to this investigation, early Fayyumic turned out to have a double system as well, with two series of definite articles, both in singular and plural.¹⁸ Table 7 below summarizes the uses of the occurrences, with the number of tokens given in parenthesis.

17 Unfortunately, in lack of native informants to test, one will never be able to decide whether this case is a real contradiction to the theory, or whether the use of the weak article might have an alternative source (e.g. the actual decision of the speaker/composer to consider the referent situationally unique and to suppress the anaphoric aspect of reference; very speculatively, a scribal error can never be excluded either). Of course, exceptions of this sort come up all the time. Dealing with a dead language will always have these limits, data can only be collected, counted and evaluated to produce statistical support for or against the theoretical claims.

18 The same survey has been carried out in another, shorter manuscript, BM. Or. 6948 (ed. by Gaselee 1909), which contains fragments from the *Acts*, and a similar system has been found as far as the forms of the articles are concerned.

Table 7 | Definite determiners in the early Fayyumic ms. BM Or. 5707

| | SG.M | SG.F | PL |
|------------------------------------|-------------------------------|-----------------------|-----------------------|
| Possessed (10) | | | nen- |
| | π-/πε- | τ-/τε- | nen- |
| Inherently/situational unique (70) | p-/pe- | t-/te- | – |
| Generic (7+7) | | | ni- |
| Direct anaphor (8) | π- pi- | τ- ti- | ni- ni- |
| Article+relative converter (16) | π-ετ- p-et- | τ-ετ- t-et- | π-ετ- n-et- |
| Article+N+relative converter (4) | π-... ετ- p-... et- | <i>no data</i> | <i>no data</i> |

The longer forms **πε-/τε-** p-/te- are attested under the same conditions as in Sahidic. Weak articles appear with possessed nouns on the one hand, and with unique nouns on the other (encoding both inherently uniqueness and situational uniqueness). Weak articles also stand with generic noun phrases in singular, while generics in plural are determined by the strong article. Strong articles can be found in direct anaphoric use elsewhere, as was expected on the basis of Bohairic patterns. It must be noted that there are two cases of strong articles which, considering the textual context, are not used anaphorically, but seem to mark contrast instead. Finally, the short, weak article regularly appears (both in singular and in plural) if it is heading a relative clause. This latter group of data seems to be special in the sense that the distinction of semantic vs. pragmatic uniqueness is neutralized in this syntactic context, and the definite determiner has a single, invariable form in this position.

6 Concluding remarks with some typological considerations

No doubt, such a complex hypothesis should be tested more widely, i.e. in a larger corpus that involves texts from more periods and from various registers (registers, for instance, might be a relevant aspect in the case of classical and late Fayyumic). In this study, only early texts have been examined, dating to the 4th or 5th centuries.

As for the moment, let me consider some typological and diachronic facts that might support the hypothesis put forward in the previous sections. The division between semantic uniqueness and pragmatic definiteness in grammatical systems that make use of two distinct sets of articles is not exceptional. A very similar distribution has been observed, for instance, in many German dialects. The two sets are usually referred to as weak vs. strong articles or reduced vs. full form of articles. Weak articles usually appear to encode (situational or larger situational) uniqueness, while strong articles are claimed to have an anaphoric nature.¹⁹

¹⁹ Languages (mainly German dialects) in which a double article system has been reported: the Frisian dialect of Fering (Ebert 1971); see also Himmelmann (1997: 54–55) with further references

Interestingly, the distinction between semantic uniqueness and pragmatic definiteness may have a crucial role in language change as well. Investigating the early use of the article in Old Hungarian, it has been claimed that the new determiner, which developed during the individual history of the language²⁰, first appeared to encode pragmatic definiteness only (Egedi 2013, 2014). The use of the article only extended to inherently unique nouns, to generics and the possessed nouns at a later stage of the language. Furthermore, in some Slavic language varieties, in which a relatively new article is in use, a similar phenomenon has been observed: the new article typically appears to mark pragmatic definiteness (cf. Czardybon (2012) for the Upper Silesian dialect of Polish; Breu (2004) and Scholze (2012) for Colloquial Upper Sorbian spoken in Eastern Saxony).

Löbner (2011), whose semantic model was adopted in this paper, also assumed that in languages where article derives from a demonstrative, it is a usual scenario that there is a stage when only “incongruent” (i.e. pragmatic) definiteness is marked by the definite article. However, this generalization has not been justified so far by a comparative and exhaustive empirical study. What has been found in the history of Hungarian suggests that distinguishing two kinds of definiteness might be revealing not only in understanding synchronic systems with two sets of articles, but in reconstructing diachronic processes as well. What is needed, therefore, is collecting and studying relevant data from languages that do have historical records, and the article emerged during their written history.

Bibliography

- Abbott, Barbara. 2004. Definiteness and Indefiniteness, in: Laurence R. Horn & Gregory Ward (eds.), *The Handbook of Pragmatics*, Oxford: Blackwell Publishing, 122–149.
- Breu, Walter. 2004. Der definite Artikel in der obersorbischen Umgangssprache, in: Marion Krause & Christian Sappok (eds.), *Slavistische Linguistik 2002*, Munich: Otto Sagner, 9–57.
- Černý, Jaroslav & Sarah I. Groll. 1978. *A Late Egyptian Grammar*, 2nd ed. Rome: Biblical Institute Press.
- Crum, Walter E. & Frederic G. Kenyon. 1900. Two Chapters of Saint John in Greek and Middle Egyptian, in: *Journal of Theological Studies* 1, 415–433.
- Czardybon, Adrian. 2012. The definite article in the Upper Silesian dialect of Polish, Talk given at the Workshop on “Semantic and Typological Perspectives on Definites”, Düsseldorf, Germany (01.06.2012)
- Depuydt, Leo. 1985. The double definite article in Bohairic, in: *Enchoria* 13, 51–63.
- Ebert, Karen H. 1971. *Referenz, Sprechsituation und die bestimmten Artikel in einem Nordfriesischen Dialekt (Fering)*, PhD thesis, Christian-Albrechts-Universität zu Kiel.

to earlier literature on Rhineland dialects; about the Low German dialects of the North Sea region and Scandinavian languages, see Schroeder (2006); about the semantics of contracted vs. non-contracted forms in Standard German prepositional phrases, see Schwarz (2009); for Swiss German, consult Studler (2011). In some languages the opposition emerges in the form of simple marking vs. double marking (e.g. Swedish). For similar double systems in other, unrelated languages, e.g. Hausa and Lakota, see Lyons (1999: 53–54).

- 20 The closest relative languages, Khanty and Mansi, do not have an article. Neither does the majority of Uralic languages have one. Whether the emergence of the Hungarian definite article is due to an internal innovation, or is the result of some sort of contact situation is still a matter of debate.

- Egedi, Barbara. 2012. *Coptic noun phrases*, PhD thesis, Budapest, Eötvös Loránd University.
- 2013. Grammatical encoding of referentiality in the history of Hungarian, in: Anna Giacalone-Ramat, Caterina Mauri & Piera Molinelli (eds.), *Synchrony and Diachrony: a Dynamic Interface*, Amsterdam: John Benjamins, 367–390.
- 2014. The DP-cycle in Hungarian and the functional extension of the noun phrase, in: Katalin É. Kiss (ed.), *The Evolution of Functional Left Peripheries in Hungarian Syntax*, Oxford Studies in Diachronic and Historical Linguistics 11, Oxford: Oxford University Press, 56–82.
- *forthcoming*. Dialectal variation in Coptic possessive constructions, in: Eitan Grossman & Stéphane Polis (eds.), *Possession in Ancient Egyptian*, Berlin: Mouton de Gruyter.
- Erman, Adolf. 1933. *Neuägyptische Grammatik*, 2. Aufl. Leipzig.
- Funk, Wolf-Peter. 1988. Dialects wanting homes: a numerical approach to the early varieties of Coptic, in: Jacek Fisiak (ed.), *Historical Dialectology: Regional and Social*, Trends in Linguistics Studies and Monographs 37, Berlin: Mouton de Gruyter, 149–192.
- 1991. Dialects, morphology of Coptic, in: Aziz S. Atiya (ed.), *The Coptic Encyclopedia* 8, New York: Macmillan, 101–108.
- Gelderen, Elly van. 2011. *The Linguistic Cycle. Language Change and the Language Faculty*, Oxford: Oxford University Press.
- Givón, Talmy. 2001. *Syntax. An Introduction*, Amsterdam: John Benjamin.
- Himmelmann, Nikolaus P. 1997. *Deiktikon, Artikel, Nominalphrase: Zur Emergenz syntaktischer Struktur*, Linguistische Arbeiten 362, Tübingen: Niemeyer.
- 1998. Regularity in irregularity. Article use in adpositional phrases, in: *Linguistic Typology* 2, 315–353.
- 2001. Articles, in: Martin Haspelmath et al. (eds.), *Language Typology and Language Universals*, Handbücher der Sprach und Kommunikationswissenschaft 20.1, Berlin & New York: Walter de Gruyter, 831–841.
- Junge, Friedrich. 1996. *Einführung in die Grammatik des Neuägyptischen*, Wiesbaden: Harrassowitz.
- Kahle, Paul E. 1954. *Bala'izah. Coptic Texts from from Deir el-Bala'izah in Upper Egypt*, London: Oxford University Press.
- Kasser, Rodolphe. 1958. *Papyrus Bodmer III. Evangile de Jean et Genese I–IV,2 bohairique*, Corpus Scriptorum Christianorum Orientalium 177–178, Louvain.
- 1991. Dialects, grouping and major groups of, in: Aziz S. Atiya (ed.): *The Coptic Encyclopedia* 8. New York: Macmillan. 97–101.
- Kroeber, Burkhart. 1970. *Die Neuägyptizismen vor der Amarnazeit. Studien zur Entwicklung der ägyptischen Sprache vom Mittleren zum Neuen Reich*, Bamberg: Aku.
- Kupreyev, Maxim. 2014. The origins and Development of the definite article in Egyptian-Coptic, in: *Studies in Ancient Art and Civilizations* 18, 223–237.
- Layton, Bentley. 2000. *A Coptic Grammar with Chrestomathy and Glossary. Sahidic Dialect*, Wiesbaden: Harrassowitz.
- Loprieno, Antonio. 1980. Osservazioni sullo sviluppo dell'articolo in egiziano e nelle lingue semitiche, in: *Oriens Antiquus* 19, 1–27.
- 1995. *Ancient Egyptian. A linguistic Introduction*, Cambridge: Cambridge University Press.
- Löbner, Sebastian. 1985. Definites, in: *Journal of Semantics* 4, 279–326.
- 2011. Concept types and determination, in: *Journal of Semantics* 28, 279–333.
- Lyons, Christopher. 1999. *Definiteness*, Cambridge: Cambridge University Press.
- Mallon, Alexis. 1907. *Grammaire copte avec bibliographie, chrestomathie et vocabulaire*, 2. ed. Beyrouth: Imprimerie Catholique.
- Neveu, François. 1996. *La langue des Ramsès. Grammaire de néo-égyptien*, Paris: Khéops.
- Polotsky, Hans J. 1968. The 'Weak' Plural Article in Bohairic, in: *Journal of Egyptian Archaeology* 54, 243–245.

- Quecke, Hans. 1984. *Das Johannesevangelium saïdisch. Text der Handschrift PPalau Rib. Inv.-Nr. 183 mit den Varianten der Handschriften 813 und 814 der Chester Beatty Library und der Handschrift M 569*, Rome: Papyrologica Castroctaviana.
- Scholze, Lenka. 2012. On the grammaticalization of the definite article in Colloquial Upper Sorbian (CUS), in: Björn Wiemer, Bernhard Wälchli & Björn Hansen (eds.), *Grammatical Replication and Borrowability in Language Contact*, Trends in Linguistics. Studies and Monographs 242, Berlin: Mouton De Gruyter, 323–353.
- Schroeder, Christoph. 2006. Articles and article systems in some areas of Europe, in: Giuliano Bernini & Marcia L. Schwartz (eds.), *Pragmatic Organization of Discourse in the Languages of Europe*, Berlin: Mouton de Gruyter, 545–611.
- Shisha-Halevy, Ariel. 1994. Pluridimensional Oppositions: Three Case Studies in Scripture (Pentateuch) Bohairic, in: Søren Giversen, Martin Krause & Peter Nagel (eds.), *Coptology: Past, Present and Future. Studies in Honour of Rodolphe Kasser*, Orientalia Lovaniensia Analecta 61, Louvain: Peeters, 225–247.
- 2007. *Topics in Coptic Syntax. Structural Studies in the Bohairic Dialect*, Orientalia Lovaniensia Analecta 160, Louvain: Peeters.
- Steindorff, Georg. 1951. *Lehrbuch der koptischen Grammatik*, Chicago & London.
- Stern, Ludwig. 1880. *Koptische Grammatik*, Leipzig: T. O. Weigel Verlag.
- Studler, Rebekka. 2011. *Artikelparadigmen. Form, Funktion und syntaktisch-semantische Analyse von definiten Determinierern im Schweizerdeutschen*, PhD thesis, Universität Zürich.
- Thompson, Herbert. 1924. *The Gospel of St. John according to the Earliest Coptic Manuscript*, London: British School of Archaeology in Egypt, Bernard Quaritch.
- Till, Walter C. 1930. *Koptische Chrestomathie für den fayumischen Dialekt, mit grammatischer Skizze und Anmerkungen*, Vienna.
- 1961. *Koptische Grammatik (Saïdischer Dialekt)*, 2. Aufl. Leipzig: VEB Enzyklopädie Verlag.
- Vergote, Jozef. 1973. *Grammaire Copte. Ia. Introduction, phonétique et phonologie, morphologie synthématique (structure des sémantèmes). Partie synchronique*, Louvain: Peeters.
- 1983. *Grammaire Copte. IIa. Morphologie syntagmatique. Syntaxe. Partie synchronique*, Louvain: Peeters.
- Worrell, William H. 1934. *Coptic sounds*, University of Michigan Studies, Humanistic Series 26, Ann Arbor: University of Michigan Press.
- Zöllner-Engelhardt, Monika. 2016. *Sprachwandelprozesse im Ägyptischen. Eine funktional-typologische Analyse vom Alt- zum Neuägyptischen*, Ägyptologische Abhandlungen 72, Wiesbaden: Harrassowitz.