

## THE MACEDONIAN AND BALKAN SLAVIC THREE-PART ARTICLE AND THE TYPOLOGY OF DEICTIC ARTICLES

**Max Wahlström**

University of Helsinki  
max.wahlstrom@helsinki.fi

**Don Killian**

University of Helsinki  
donald.killian@helsinki.fi

Macedonian's tripartite article system (proximal, neutral, and distal) aligns with similar systems found in other Balkan Slavic varieties, but also exhibits parallels with non-Slavic languages, such as Wolof, which feature deictic definite articles. This paper investigates the morphosyntactic and semantic features of these articles in Balkan Slavic, aiming to clarify their status within the broader typology of definiteness and deixis. We show through the study of the Macedonian spoken corpus that there are significant differences between the articles regarding their relative frequency. Cross-linguistic comparison further informs the typological status of these articles. In both Balkan Slavic and Atlantic languages like Wolof, definite articles are enclitics, with one form generally occurring significantly more frequently than the others. While interesting secondary functions such as nominal tense have been proposed for these articles, these features are not diagnostic of their "articlehood" and may be expressed by demonstrative pronouns in other languages.

**Keywords:** deixis, definiteness, proximal, distal, anaphoric reference.

## МАКЕДОНСКИОТ И БАЛКАНСКО СЛОВЕНСКИОТ ТРОЕН ЧЛЕН И ТИПОЛОГИЈАТА НА ДЕИКТИЧКИТЕ ЧЛЕНОВИ

### Макс Валстром

Хелсиншки универзитет  
max.wahlstrom@helsinki.fi

### Дон Килиан

Хелсиншки универзитет  
donald.killian@helsinki.fi

Системот на тројниот член во македонскиот јазик (проксимален, неутрален и дистален) има заеднички црти со таквите системи во другите балкански словенски дијалекти, но има допирни точки и со несловенските јазици што имаат деиктички определени членови. Во овој труд ги истражуваме морфосинтаксичките и семантичките карактеристики на тројните членови во балканските словенски јазици, со цел да го разјасниме нивниот статус во пошироката типологија на определеноста и на деиксисот. Спроведената анализа на податоците од македонскиот говорен корпус покажа дека постојат значителни разлики помеѓу тие членови, во однос на нивната релативна застапеност. Споредбата со другите јазици дополнително го открива типолошкиот статус на овие членови. Во балканските словенски јазици и во атлантските јазици, како што е Волоф, определените членови се енклитики, од кои еден член значително доминира над другите во поглед на употребата. Постои мислење дека секундарните својства на овие членови не ја определуваат нивната категоријална припадност и често се изразуваат преку демонстративни заменки во другите јазици.

**Клучни зборови:** деиксис, определеност, проксимален, дистален, анафорска референција.

## 1 Introduction

The study of the Macedonian tripartite article system over the past decades has both questioned (Topolinjska 2006) and defended (Boronnikova 2014) the status of the proximal and distal markers as definite articles. Some researchers offered a more nuanced interpretation of their functions with no expressed stance on their article status (Sonnenhauser 2009). However, none of these attempts have sought to compare the Macedonian system cross-linguistically beyond other Balkan Slavic varieties, nor employed a definition of the definite article that allows to distinguish it from demonstrative pronouns. This paper seeks to contribute to the discussion by suggesting a theoretical basis for crosslinguistic comparison, briefly introducing findings from other languages with so-called deictic articles, and performing a short corpus study examining the distribution of the articles in spoken Macedonian.

Deictic definite articles are found in several unrelated languages, and are characterized by having more than one element thought to be a definite article, some of which carrying a deictic “flavor” that contrasts with a deictically neutral article. Yet many influential definitions of the definite article are based on the absence of deictic functions – this is also the premise of Topolińska’s (Topolinjska 2006) analysis, which treats the proximal and distal markers as clitic demonstrative pronouns. Moreover, the grammaticalization of demonstrative pronouns into articles is typically described as involving a loss of deixis (see, e.g., Lyons 1999: 331–332).<sup>1</sup>

These seemingly contradictory perspectives raise two primary questions: How do the three articles, especially the proximal and distal, differ from deictic demonstrative pronouns, and how does their distribution compare cross-linguistically? Additionally, what added functions do the deictically marked articles have? In this paper, we propose adhering to a typologically tested definition of definiteness that allows us to organize the observations and impressions regarding the deictic article systems. Second, we discuss non-Slavic languages, including Wolof (Atlantic), Ambel (Austronesian), and Classical Armenian (Indo-European), that have been said to display deictic definite articles. Third, we perform a small corpus study on the Macedonian spoken corpus (Escher and Winistörfer [eds.] 2021) to extract basic findings regarding the relative frequency of the three articles, and how they combine with other determiners such as possessive and demonstrative pronouns. The Macedonian findings are then discussed together with what has been reported regarding other Balkan Slavic varieties. In the conclusion, we seek to contextualize Macedonian among other languages claimed to have deictic definite articles, and propose further avenues for research.

## 2 Defining definite articles

This paper relies on a recent typological study by Laura Becker (2021) on definite articles that summarizes much of the debate of the past decades and establishes concise definitions, informed by an exhaustive cross-linguistic study of article systems in the languages of the world. According to Becker (2021: 86): “A definite article is an article that systematically marks anaphoric, recognitional, establishing, situationally unique, contextually unique, and bridging referents. It may also occur in other types of definite or generic contexts. It does not mark a referent as specific or non-specific”.

Let us now look at the six defining contexts closer.

- *Anaphoric* reference recalls something mentioned earlier in the conversation: *Leon found a book on the table. The book was surprisingly old, given its condition.*

---

<sup>1</sup> Yet Lyons specifically mentions Macedonian when stating that some articles have not lost deictic distinctions.

- *Recognitional* use contrasts with anaphora: the article signals that both the speaker and the hearer recognize the referent through shared experience or knowledge, but there is no previous mention:  
*We should go to the park this weekend.* (The speaker assumes the listener knows which park is being referred to, because it is already familiar.)
- *Establishing* referents are not identifiable to the hearer at the time of uttering; the speaker signals to the hearer that they can or should construct the referent as (soon to be) identifiable:  
*Did you hear the news? Some otters escaped from the zoo.*
- *Situationally unique* referents are identifiable by being the only available ones in the discourse context:  
*Please write your answer on the blackboard.*
- *Contextually unique* referents, on the other hand, are identifiable in a larger context in which they are unique:  
*The prime minister (of the country we live in) held a preference conference today about the security leak.*
- *Bridging referents* are identifiable through a link that exists between something that has been said before in the discourse; Becker (2021: 81) distinguishes between two different types of bridging referents:
  - *Unique bridging referents* are identifiable by both the speaker and the hearer through an unambiguous link such as a part–whole relationship:  
*I was driving my bike when the rear wheel suddenly went flat.*
  - *Relational bridging referents* in contrast display a looser, less direct link, but the speaker signals that the referent is identifiable in relation to a previously mentioned referent:  
*We tried out this new restaurant yesterday, and the food was excellent.*

While some of these defining contexts are shared with demonstratives, Becker (2021: 103) goes further and identifies contexts in which only definite articles may occur. According to her, only definite articles encode situationally and contextually unique referents, and unique bridging referents. The distinction between unique and relational bridging referents may seem subtle, but modern spoken Finnish offers evidence that this is a real cut-off point between demonstratives and articles. Finnish has often been discussed in the context of developing grammatical marking of definiteness (e.g., Laury 1997). Crucially, the candidate incipient article, the deictically neutral demonstrative pronoun *se*, cannot be used to mark unique bridging referents (example 1), whereas it can be used with relational bridging referents (example 2).<sup>2</sup>

- (1)            *mä        ajoin    pyörällä ja ?se takapyörä hajos*  
                  I        drove bike        and DEM rear\_wheel broke  
                  ‘I was driving a bike and the rear wheel\_broke.’  
                  (Finnish, own knowledge)

<sup>2</sup> Glosses follow Leipzig standards except for the following: AOR Aorist, ART Article, CLSB Class B, EST Established, FOC Focus, I Inclusive, INSTR Instrumental, INT Interrogative, MID Mid-distance, NHUM Non-Human, NVIS Non-visible, ORI Orientative, PRO Pronoun, REC Recipient, REF Referential, REM Remote Deixis, REMPST Remote Past, TR Transitive, VIS Visible.

- (2) *käytiin yhes ravintolas ja se ruoka oli hyvää*  
 we\_went one restaurant and DEM food was good  
 ‘we went to a restaurant and the food was good.’  
 (Finnish, own knowledge)

What this definition entails, is that in order to consider the Macedonian proximal and distal elements as definite articles, they must occur in contexts in which demonstratives cannot, as suggested also by Topolińska (Topolinjska 2006). In what follows, we observe a few parallel systems featuring so-called deictic articles.

### 3 Towards the typology of deictic definite articles

Deictic definite articles follow the same definition outlined previously; that is, they occur in “anaphoric, recognitional, establishing, situationally unique, contextually unique, and bridging contexts.” However, they also include a deictic component in their meaning, the precise function of which varies from language to language. Lyons (1999: 55) states that “The kind of deictic distinctions which typically occur in demonstratives (distinctions of distance from the speaker, association with different persons, for example) are occasionally found in simple definites.” He further observes that proximity features appear independently of demonstratives, and that, in some languages, the same deictic features are present on both demonstratives and the definite article. Yet, as the data below show, the situation is more complex than Lyons describes.

The referent of a deictic definite article does not have to refer to an object present in the immediate discourse situation; Becker states that “deictic referents can but do not need to be marked by definite articles” (2021: 93). Instead, the languages claimed to have deictic definite articles divide these articles into different categories based on spatial deixis. Examples of such languages include deictic articles in the Atlantic languages Saafi-Saafi, Wolof and Noon, the Austronesian languages Nemi and Ambel, South Slavic Rhodopian, Torlak varieties, and Macedonian.

Wolof, for example, exhibits a two-way deictic distinction in its definite article system, with proximal *Ci* and distal *Ca*, *C* indicating the noun class of the referent the article modifies. Becher (2001:65) contrasts Wolof definite articles using *buur* (king) as an example:

- *buur bi*: the king here; the king who currently reigns
- *buur ba*: the king there; the king of another country or of times gone by

Although the examples Becher gives are of archaic Wolof, Modern Dakar Wolof functions similarly with respect to *Ci* and *Ca*. However, despite clear morphological connections to the demonstratives, distributional analyses suggest that the definite articles do not operate in the same way as demonstratives. The proximal is far more common than the distal, and is also used when deixis not considered relevant for the speaker. This analysis has also been adopted by researchers specializing in Atlantic.<sup>3</sup>

In a small corpus of 5 texts of Wolof (Robert 2015), out of 170 total definite articles, 147 were marked with proximal and 23 with distal. The distal examples are typically used to indicate clear distance in time or space, and referent tracking does not seem to be relevant to the distinction.

<sup>3</sup> Denis Creissels (p.c.) mentions that *Ci* is clearly not proximal, but rather default, contrary to the demonstrative *Cii* from which it originates.

- (3) *waaw* *ñoom* *ñoo* *sanc* *dëkk* *b-a*  
 yes PRO.3PL FOC.SBJ.3PL implant village CLSB.DET-DIST  
 ‘Yes, they're the ones who founded the village.’  
 (Robert 2015: WOL\_SR\_CONV\_01\_SADDA)
- (3) *te* *mu* *ne-aan* *sama* *naar* *b-a*  
 and AOR.3SG say-REMPST POSS.1SG moor CLSB.DET-DIST  
*des* *c-a* *Aalëyétu*  
 remain LOC-DIST Aalëyétu  
 ‘and about whom he said, "my Moor who remained at Aalayetu.”’  
 (Robert 2015: WOL\_SR\_NARR\_04\_FALLU-STICK)
- (4) *moo-y* *y-a* *maam* *y-a* *yor-oon*  
 FOC.SBJ.3SG-IPFV CLSY.PL-DIST grandparent CLSY.PL-DIST hold-PST  
*te* *ñu* *sàggan-e=ko*  
 and AOR.3PL be\_careless-APPL.NHUM=OBJ.3SG  
 ‘is that, what the ancestors held and which has been neglected.’  
 (Robert 2015: WOL\_SR\_NARR\_03\_FALLU.EAF)

All of the languages listed above except for Ambel have similar article systems in that spatial deixis is obligatory across the entire system. While we do not have sufficient data for most of these languages to confirm this definitely, we predict that each language of this type also has some kind of “default” form for when deixis is not relevant.

Ambel has a considerably more complex system, with two types of non-spatial definite articles based on accessibility to the hearer, 32 different spatially-oriented deictic definite articles, and three indefinite articles that differentiate specificity.

- (5) *mokoné: 'nya-kabút* *kalamlí* *lu-pa* *be* *tuta-la*  
 say.3SG 2SG-*hold* scoop *sea-MID* and 1DU.I-ORI  
*líl* *tuta-mát* *mi* *láp* *lí-ma'*  
 land 1DU.I-*die* INSTR fire *land-DIST*  
 ‘He said: “Grab the (seawards) scoop and let's go landwards to extinguish the (landwards) fire.”’  
 (Laura Arnold, p.c.)

In addition to definite deictic articles, languages such as Musqueam Halkomelem (Salishan) may incorporate spatial deixis as a component in their article system, but with a broader range of reference types including indefinite (specific) and indefinite (nonspecific) reference. Becker (2021) classifies Musqueam Halkomelem as having “referential articles,” split into three deictic categories: proximal and visible, proximal and non-visible, and remote.

- (6) *niʔ* *skʷtəxʷ* *ʔə* *tə* *léləm* *kʷθə* *sqʷəməy*  
 be.there inside OBL ART:REF.VIS house ART:REF.NVIS dog  
 ‘The dog is in the house.’  
 (Suttles 2004: 342)
- (7) *ʔi* *ʔəmí* *ʔáʔqəl* *tə* *sqʷəməy*  
 be.here come exit ART:REF.VIS dog  
 ‘The dog came out.’  
 (Suttles 2004: 342)

These articles also occur with indefinite semantics (hence the label of referential rather than definite). Nonspecific indefinite referents are generally marked by the remote form of the article, while specific indefinite referents are marked by either the non-visible or visible articles.

- (8) *s-ni<sup>2</sup>-ct*                      *wə-k<sup>w</sup>ác-nəx<sup>w</sup>*                      *tə*                      *čiči<sup>2</sup>q<sup>2</sup> ən*  
 NMLZ-AUX-our                      EST-see-TR                      ART:REF.VIS                      mink  
 ‘Then we saw a mink.’  
 (Suttles 2004: 347)
- (9) *ʔéx<sup>w</sup>-əs-t-Samx*    *čx<sup>w</sup>*                      *ʔə*                      *k<sup>w</sup>ə*                      *léctən*  
 give-REC-TR-me you                      OBL                      ART:REF.REM                      knife  
 ‘Give me a knife.’  
 (Suttles 2004: 51)

Classical Armenian is also said to have deictic definite articles. Classical Armenian features a three-part system of proximal (-s-), medial (-d-), and distal (-n-) deixis (Müth 2011: 12–13). These elements occur, in addition to, for instance, demonstrative pronouns, as enclitics affixed to nouns, and are considered in the literature definite articles denoting personal affinity: -s refers to the speaker, -d to the addressee, and -n as a neutral definite article. According to Müth (2011: 14–15), Classical Armenian, unlike Greek, generally avoids definite articles in expressing “semantic definiteness,” such as with proper nouns, unique reference nouns, generic reference nouns, and nouns determined by superlative, comparative, or ordinal attributes. Importantly, Müth offers examples of the Classical Armenian article use in “pragmatic” contexts that include the marking of anaphoric, establishing, and both bridging and unique referents. However, in all these examples, only the distal article -n is employed; the study thus does not provide evidence of the proximal and medial elements being used as definite articles.<sup>4</sup>

Finally, another language with proposed deictic definite articles, often discussed in connection with Armenian due to a long-standing areal connection, is Common Kartvelian. According to Harris (1985: 75–77), in what the author refers to as Oldest Georgian, there existed a system of three definite articles: “proximate,” expressing closeness to the speaker, “contingent,” close to the addressee, and “remote,” distanced from both discussants. Hodgson (2022: 128) suggests that prehistoric Kartvelian may have played a role in the development of the Classical Armenian system.

#### 4 Macedonian and Balkan Slavic

In this section, we summarize findings regarding Balkan Slavic. We discuss first what has been argued in the previous literature. We then observe the distribution of articles in a Macedonian speech corpus, and briefly discuss the key features of the deictic articles in other South Slavic varieties such as Torlak and Rhodopian.

Topolińska (Topolinjska 2006: 9–10), who does not classify the Macedonian proximal and distal elements as definite articles, contrasts the neutral definite article with the full

<sup>4</sup> The status of the definite article that no longer displays deictic distinctions is complicated in Modern Armenian as well. Zolyan (2024) speaks of incomplete grammaticalization of the Modern Eastern Armenian definite article, which retains some of its original demonstrative and possessive meanings. The definite article is used in the nominative and accusative cases to mark subjects and direct objects, and its use is not strictly tied to the semantic characteristics of definiteness or indefiniteness but is more related to syntactic positions. Some argue it is used to mark specific referents (see Hodgson 2022: 146).

demonstrative pronoun containing the same element *-t-*. According to her, using the demonstrative pronoun with situationally and contextually unique referents is either not possible, or results in a change of meaning. Moreover, she argues that the neutral definite article is the only unmarked device that conveys anaphoric reference.

Another context in which the proximal and distal elements do not occur is the expression of generic referents (Usikova 2000: 79).<sup>5</sup> In Becker's (2021: 86) typology, however, definite articles optionally mark generic reference in some languages, but it is not a defining characteristic.

Topolińska (Topolinjska 2006: 11–13) proceeds to analyze the particular contexts for the proximal and distal elements. These include the expression of spatial deixis functioning similarly to demonstrative pronouns. The contexts in which one could use the neutral *-t-* may further convey a sense of familiarity or closeness to the speaker with the proximal, or temporal and emotional distance with the distal. These functions have been further elaborated by Sonnenhauser (2009), who emphasizes the subtle nuances they introduce, particularly in constructions involving coordination of anaphoric reference. Boronnikova (2014), who defends the status of all three clitic elements as definite articles, highlights their expressive value, describing the proximal as carrying a positive tone, and the distal a negative one.

Both Boronnikova and Topolińska agree that the status of the proximal and distal clitic elements hinges upon their relationship with the demonstrative pronouns. Topolińska, on the one hand, demonstrates how the neutral clitic element differs from the corresponding neutral demonstrative pronoun, carrying certain characteristics that are exclusive to definite articles. However, she argues that the proximal and distal forms reflect the same deictic distinctions as the full demonstrative pronouns and thus do not qualify as definite articles. Boronnikova, in contrast, is less interested in defining the grammatical criteria of definiteness and instead challenges Topolińska's claim about the functional equivalence of the clitics and pronouns. Boronnikova (2014: 63–65) highlights the fact that in spoken Macedonian the proximal and distal clitics may occur in the same NP, which, according to her, proves that they do not serve the same function as demonstratives.<sup>6</sup>

#### 4.1 Corpus study of spoken Macedonian

This small corpus study utilizes two subcorpora of the Macedonian Spoken Corpus (Escher and Winistörfer 2021): The first subcorpus consists of field data from Western Macedonian dialects (in this paper we call the subcorpus *Dialects*), collected from informants in Resen, Janche, Krani, and Arvati, with 90% of the speakers born between 1948 and 1956. In this paper, we only use data from speakers with Macedonian as their first language, resulting in a dataset of 15 informants and 18,542 tokens. The second subcorpus, called *Bombi*, consists of 141,990 tokens and represents the modern urban variety of Skopje, including traits from various

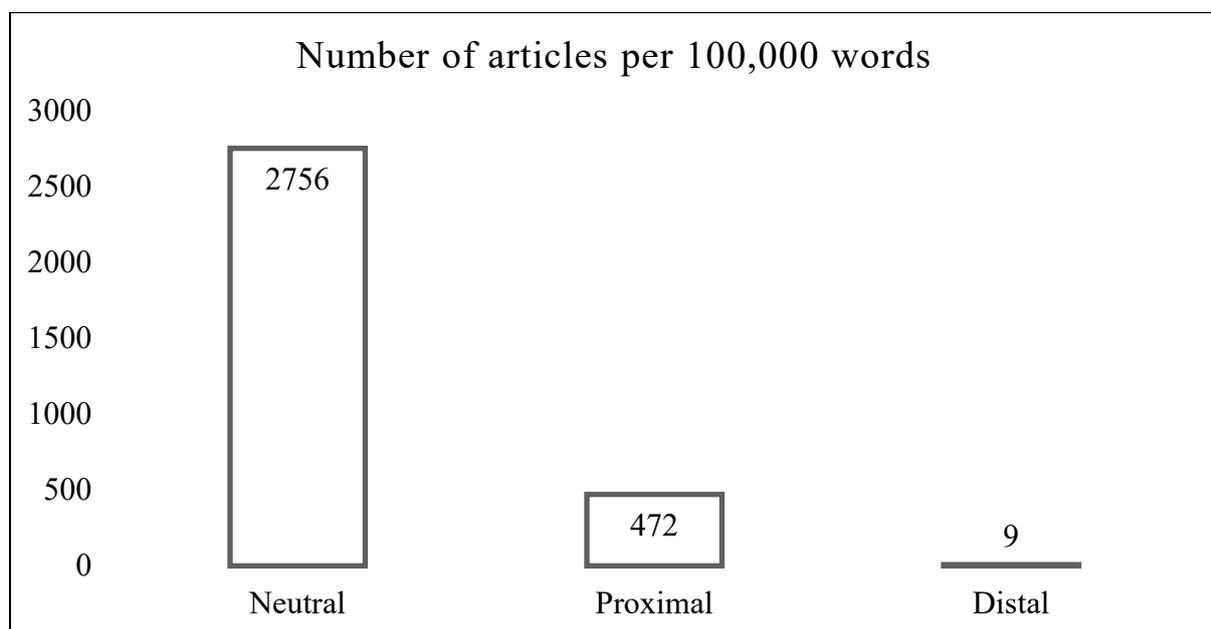
<sup>5</sup> See also Karapejovski (2022) for recent study on genericity and definiteness in Macedonian.

<sup>6</sup> Additionally, Boronnikova (2014: 64) gives an example of “generic” use of the distal element: *I ko ke beše blinkerot, trebaše četiri ribi da se fatet, tie fašćea po deset, petnaeset, koj kako, bez kontrola. Ribar-on da fatit deset kila, tri kila ke odnesit na ribarnica i od kaj znajt ovoj, nemat financova kontrola.* ‘And when they were allowed to catch four fish with a lure, they caught ten to fifteen kilograms, as much as they could, without control. The/a/that fisherman catches ten kilos but only takes three to the fishmonger, and how can they know; there is no financial control.’ It is true that the referent is non-specific, but even non-specific referents can be sometimes picked up in discourse; these are what Karttunen (1976) calls short-term discourse referents. In this case, what the highlighted noun represents is rather a relational bridging referent (see Example 2). Although this function can also be conveyed by a demonstrative pronoun, this is particularly interesting because, as Boronnikova points out, the distal element seems to convey a disdainful attitude towards the described practice.

regional dialects due to migration to the capital. There are more than a hundred speakers, with men being overrepresented by 5 to 1.<sup>7</sup>

Research resources for modern spoken Macedonian remain very limited, and this corpus represents therefore an incredibly valuable trailblazing endeavor. However, due to the preliminary nature of the corpus, the accuracy of the automated POS tagging, lemmatization, and morphological parsing is not sufficiently high to replace the search of word forms.<sup>8</sup> We therefore restrict ourselves to establishing only some of the more basic characteristics of the articles in the data along with a few additional observations about their distribution. This topic undoubtedly merits a more thorough analysis, but due to time constraints, that will have to wait for a later opportunity.

We first sought to study the relative frequencies of the three articles. The three articles of Macedonian belong to entirely different frequency categories, as shown in Figure 1. The distal one is extremely rare, the proximal is more than 50 times more common, yet still only one-sixth as frequent as the neutral article.

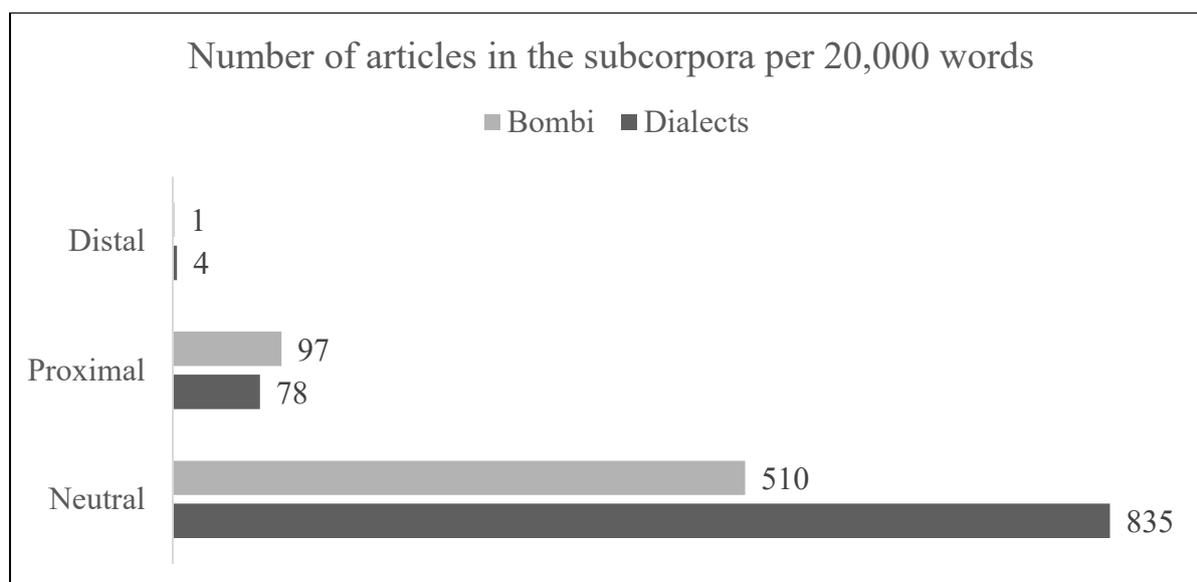


**Figure 1.** Frequency of the articles in the *Dialects* and *Bombi* subcorpora of the Macedonian Spoken Corpus (Escher and Winistörfer 2021), n=160,532

The corpus provides some opportunities to assess the homogeneity of the data through its metadata. Figure 2 illustrates the distribution of articles across the two subcorpora. The frequency of the proximal article is consistent across the subcorpora; however, the occurrences of the distal article are too few to allow for a reliable determination. Overall, the frequency scale for all articles remains relatively stable.

<sup>7</sup> The subcorpus contains transcripts of wiretapped conversations of Macedonian political elites from 2008 to 2015 in modern colloquial Macedonian used by educated elites. These transcripts were published by the opposition party SDSM in 2015, revealing the government's high level of corruption. (Friedman 2017.)

<sup>8</sup> Also, at the time of writing this mid-March 2025, the home page of the corpus is offline.



**Figure 2.** Frequency of the articles per two subcorpora of the Macedonian Spoken Corpus (Escher and Winistörfer 2021), n=160,532

The *Dialects* subcorpus further enables the evaluation of variation related to specific interviews or speakers, although there are enough occurrences only for the neutral article. The frequency of the neutral article in the *Bombi* subcorpus is 2.6%, while the frequencies among speakers in the *Dialects* subcorpus range from 3.1% to 6.9%. This high variance is likely not indicative of permanent linguistic idiosyncrasies of the speakers, but rather is more likely due to chance factors, such as the topics of conversation. Notably, the shortest inputs from a single speaker consist of fewer than 500 words in total.

The variance is a crucial indicator of the sufficiency of corpus size in assessing the definite article. With a corpus of fewer than 200,000 tokens, determining the relative frequencies of the three articles is barely feasible. However, obtaining a representative sample of contexts for especially the distal article, given its low frequency, would require a corpus at least ten times larger.

For the reasons outlined above, the lack of diagnostic contexts for definite articles regarding the proximal and distal clitic elements in this corpus does not serve as evidence against their articlehood. The corpus is simply too small to definitively assess this. However, the low frequency of the proximal as well as the distal elements in particular raises questions about one general condition in Becker's (2021: 86) definition: a definite article must *systematically* mark referents in the six defining contexts. With certainty, we can only confirm that the neutral article fulfils this condition.

## 4.2 Torlak and Rhodopian

We now turn briefly at the closely related South Slavic varieties of Torlak, spoken in Southeastern Serbia, as well as the Rhodopian dialects of Bulgaria, including Xanthi Pomak, which is part of the same macrodialect.

Vuković (2023) examines clitic demonstrative elements, traditionally classified as definite articles in Torlak, also known as Prizren-Timok, a transitional linguistic variety or a set of varieties between Macedonian and Bulgarian on the one hand and Serbian on the other. The variety Vuković focuses on in particular is Timok Torlak, also known as Prizren-Timok. The *-v-*, *-t-*, and *-n-* elements in Timok Torlak are cognates with the Macedonian ones discussed above. According to Vuković (2023: 265–266), the Timok Torlak demonstrative clitics are not

full definite articles, but, at least, in the case of the neutral *-t-* clitic, they are reminiscent of an anaphoric article. The proximal and distal elements, on the other hand, are more commonly used deictically.

The relative frequencies Vuković (2023: 254) presents are interesting when compared to spoken Macedonian: there are 147 *t*-stem (neutral), 26 *v*-stem (proximal), and 4 *n*-stem (distal) clitics per 10,000 nouns. Since the compared items are nouns, rather than words as in Figures 1 and 2, these results are not directly comparable regarding absolute frequencies. However, assuming that nouns make up no more than half of all the words in the corpus, the neutral clitic is at least three times rarer than its Macedonian counterpart. This provides giving frequency-based support for Vuković's conclusion that Timok Torlak does not have a definite article.

The Rhodopian dialect of Bulgarian, along with closely related Pomak spoken by Muslim Slavs in Northern Greece, both display a three-part system of postposed clitic elements. The key sources regarding the Rhodopian three-part system are Kanevska-Nikolova's (2006) monograph, along with Fanciullo's (2019) in depth analysis. The two works together offer a wealth of discussions and examples, much beyond the scope of this paper. The morphological elements of the three-part system is similar to those of Macedonian and Torlak, but the proximal element derives from a different Slavic demonstrative stem, *-s-*, not *-v-*. However, in other respects Rhodopian differs significantly from the other Balkan Slavic three-part systems: both authors present examples involving other elements of contexts that suggest a much more complicated situation.

According to Fanciullo (2019), the choice of elements depends on spatial interpretations, evidential values, temporal relations, the degree of familiarity, the speaker's attitude toward the information, the word class of the referent, and whether the referent is an inalienable property of the speaker.

Kinship nouns, for instance, can associate the deictic elements with spatial meanings, but these elements can also indicate the relationship of the members (close relationships associated with the *-s-* element, and distant kinship relations with *-n-*). In example (10), *-s-* indicates physical proximity (come *here* often), whereas *-s-* in (11) indicates both a close degree of kinship and possession.

- (10) Rhodope dialect  
*'Dɔftera-sa ti du'hoda li si 'fostif?*  
 daughter-DEF.S POSS.2SG come.PRS.3SG INT REFL often  
 'Does your daughter come home often?'  
 (Kanevska-Nikolova 2006: 68, cited in Fanciullo 2019: 58)

- (11) Rhodope dialect  
*ja 'ima-m si gra'dinka, ra'zdeli-me*  
 1SG have.PRS-1SG REFL garden divide.PRS-1PL  
*i na sna'ho-sa*  
 and to daughter-in-law-DEF.S  
 'I have a vegetable garden, we divide [everything] to (my) daughter-in-law...'  
 (Fanciullo 2019: 109)

Definite deictic articles can also express nominal tense. In the following examples, the speaker explains the choice of a particular deictic element, that *-s-* refers to the present tense.

- (12) Rhodope dialect  
*zaštò kàzva glàdni-se i ne kàzva*  
 why say.PRS.3SG hungry-DEF.S and NEG say.PRS.3SG  
*glàdni-te*  
 hungry-DEF.T  
 ‘Why does he say the[s] hungry ones and not the[t] hungry ones?’  
 (Fanciullo 2019: 153–154)
- (13) Rhodope dialect  
*glàdni-se zàštòtu f mum'èntə sə glàdni*  
 hungry-DEF.S because in moment be.PRS.3PL hungry  
 ‘The[s] hungry ones, because it’s right now that they are hungry!’  
 (Fanciullo 2019: 153–154)

This expression of temporal values is also possible with unique referents, as in example (14).

- (14) 'mnogo pe'tʃe 'slontse-so  
 very shine.PRS.3SG sun-DEF.S  
 ‘The sun is shining a lot.’  
 (Fanciullo 2019: 131)

According to Fanciullo's (2019) corpus study of Rhodopian, the total amount of deictic *-t-* in the analyzed corpora was 899 (relative frequency 77.5 %), the deictic *-n-* occurred 170 times (relative frequency 14.6 %), and the total number of the deictic *-s-* was 90 (relative frequency 7.7 %). However, assessing relative or absolute frequencies of the elements is complicated by a number of confounding factors, such as the influence of standard Bulgarian, which has only one clitic originating in a demonstrative pronoun, the definite article *-ăt*. Moreover, the use of the elements varies depending on the age and location of the speakers; the use of the element *-t-* is relatively constant, but *-s-* is mainly used by older speakers, and its frequency in the corpus is proportional to the age of the speaker (Fanciullo 2019). The element *-n-* is used by all age groups, but is also more frequent among elderly speakers. Geographically, the tripartite system of deictics is best represented near Smolyan; *-s-* elements in particular are more attested in villages near the center. The frequency of use of *-s-* decreases proportionally with distance from Smolyan (Fanciullo 2019).

## 5 Conclusions

This article has aimed to contribute to the understanding of the Macedonian so-called tripartite article in several ways: by comparing it with other similar systems in the languages of the world known as deictic articles, by observing frequency-related characteristics of the articles in spoken Macedonian, and by summarizing the previous discussion on the Balkan Slavic systems of deictic articles. Crucially, through introducing a clear-cut definition of the definite article, we have sought to provide answers to the open questions regarding their status.

Regardless of whether all three elements of the Macedonian tripartite article should merit the status of definite article, Macedonian clearly belongs to a very small global group of languages that have what are known as deictic articles. In fact, based on our limited typological sample, with one exception, deictic articles seem to have a default form for when deixis is not

relevant – and that this is likely to be reflected in the relative frequencies of the different articles, with one article dominating over others.

What this entails for the less common articles remains unclear. It seems that the nomenclature defining all the contributing elements to be definite articles is based primarily on morphosyntactic observations: the language in question clearly has grammaticalized marking of definiteness, and all complementarily distributed morphosyntactically and semantically similar elements are called articles, often without much consideration for contexts that differentiate definite articles from other linguistic elements. Therefore, the debate surrounding articlehood of the Macedonian proximal and deictic clitics may mirror the debate faced by other languages with deictic articles, which are, by default, less studied and more poorly resourced than Macedonian.

The debate surrounding the Macedonian language has already brought to light most of the crucial aspects regarding the status of the tripartite article. It appears that there is little disagreement about the empirical facts themselves, but rather about how these elements should be categorized and named. However, without a solid and widely accepted definition of the definite article, it is impossible to resolve the key question. Topolińska (Topolinjska 2006) rightly suggests that only the neutral clitic consistently appears in contexts typically associated with definite articles. Our observations, along with those concerning the Torlak, regarding the relative frequencies of the elements support the argument for recognizing a single definite article.

Boronnikova (2014), on the other hand, makes a valuable point by arguing that clitics and full demonstrative pronouns should not be treated as synonymous. Additionally, observations from Macedonian (Sonnenhauser 2009), Rhodopian and Xanthi Pomak (Fanciullo 2019; Adamou 2011), and Wolof (Becher 2001) suggest that a functional divergence from demonstrative pronouns is indeed characteristic of deictic articles.

Phenomena related to the grammatical or grammaticalized marking of definiteness require extensive data, as we have demonstrated. It is therefore of utmost importance that larger speech corpora of modern spoken Macedonian become available, especially given the likelihood of significant areal variation.<sup>9</sup> Other methods, such as grammaticality judgments, can be used to determine whether the proximal and distal elements occur in diagnostic contexts of situationally and contextually unique referents and unique bridging referents. However, to fully understand the Macedonian deictic article, it must be ultimately examined within a large, annotated corpus.

## Bibliography

- Adamou, E. (2011). Temporal uses of definite articles and demonstratives in Pomak (Slavic, Greece). *Lingua*, 121(5): 871–889. DOI: 10.1016/j.lingua.2010.12.001.
- Becher, J. (1991). *Untersuchungen zu den Verbalweiterungen des Wolof*. MA thesis. Universität Hamburg.
- Becker, L. (2021). *Articles in the World's Languages*. Berlin/ Boston: de Gruyter. DOI: 10.1515/9783110724424.
- Boronnikova, N. V. (2014). Status trojnogo chlena v makedonskom jazyke. *Filologicheskie nauki. Voprosy teorii i praktiki*, 10(3): 60–65.
- Cissé, M. (2007). *Dictionnaire Français-Wolof*. Paris: L'Asiathèque, Dictionnaire des Langues O'.
- Diessel, H. (2012). Deixis and demonstratives. *An International Handbook of Natural Language Meaning*, 3: 2407–2431. Berlin: de Gruyter. DOI: 10.1515/9783110724424.

<sup>9</sup> We would like to thank an anonymous reviewer for reminding us about the dialectal variation, specifically regarding the Ohrid dialect, an example of which is cited in this paper in the footnote on Boronnikova 2014.

- Escher, A. and Winistörfer, O. (eds.). (2021). *Macedonian Spoken Corpus*. Zürich: UZH Institute of Slavic Studies. [escher.pythonanywhere.com](http://escher.pythonanywhere.com), 11 April 2024.
- Fanciullo, D. (2019). *Temporal expression in nominals: tripartite deictics in the Bulgarian Rhodope dialects*. Berlin: Peter Lang. DOI: 10.3726/b15767.
- Friedman, V. (2017). Self and other (svoj i chuzhoj) in the Macedonian ‘bombs’ of 2015. In M. Makartsev and I. Sedakova (eds.), *Balkanskij tezaurus: Vzgljad na Balkany izvne i iznutri. Balkanske Chtenija 14, Moscow, 19–21 April 2017*, 216–229. Moscow: Institut slavjanovedenija, RAN.
- Hanks, W. F. (1992). The indexical ground of deictic reference. In A. Duranti and C. Goodwin (eds.), *Rethinking context*. Cambridge: Cambridge University Press.
- Harris, A. C. (1985). *Diachronic syntax: the Kartvelian case*. Orlando: Academic Press.
- Hodgson, K. (2022). Grammaticalization of the definite article in Armenian. *Armeniaca. International Journal of Armenian Studies*, 1: 125–150.
- Kanevska-Nikolova, E. (2006). *Trojnoto chlenuvane v rodopskite govori*. Plovdiv: Universitetsko izdatelstvo “Paisij Hilendarski.”
- Karapejovski, B. (2022). Razvojot na postpozitivniot morfološki člen vo makedonskiot jazik kako eksponent na generičnata referencija. *Slovene*, 11(2): 244–263. DOI: 10.31168/2305-6754.2022.11.2.11
- Karttunen, L. (1976). Discourse referents. In J. McCawley (ed.), *Syntax and Semantics, 2: Notes from the linguistic underground*, 363–385. New York: Academic Press.
- Laury, R. (1997). *Demonstratives in interaction: The emergence of a definite article in Finnish*. Amsterdam: John Benjamins.
- Lyons, C. (1999). *Definiteness*. Cambridge: Cambridge University Press.
- Müth, A. (2011). Categories of definiteness in Classical Armenian. *Oslo Studies in Language*, 3(3): 11–25.
- Robert, S. (2015). *Recorded and annotated oral corpus of Wolof*. CorpOrAn: Corpus Oraux Annotés, [https://corporan.huma-num.fr/Archives/corpus.php?codeLangue=code\\_WOL](https://corporan.huma-num.fr/Archives/corpus.php?codeLangue=code_WOL), 12 August 2024.
- Sonnenhauser, B. (2009). The Macedonian tripartite article: a discourse-oriented account. *Makedonski jazik*, 60: 123–136.
- Usikova, R. (2000). *Makedonskij jazyk. Grammaticeskij ocherk, teksty dlja chtenija s kommentarijami i slovarem*. Skopje: Filološki fakultet “Blaže Koneski.”
- Vuković, T. (2023). A corpus-based analysis of the grammatical status of short demonstratives in the Timok dialect. *Journal of Slavic Linguistics*, 31(1–2): 245–270.
- Topolinjska, Z. (2006). Trojniot člen: Da ili ne? *Južnoslovenski filolog*, 62: 7–15.
- Zolyan, S. (2024). On the so-called “definite article” in Eastern Armenian: Grammatical constraints and pragma-semantic functions. *Indoevropskoe jazykoznanie i klassičeskaja filologija*, 28: 641–672.