

Chapter 1

Nominal anchoring: Introduction

Kata Balogh
Heinrich-Heine-Universität Düsseldorf

Anja Latrouite
Heinrich-Heine-Universität Düsseldorf

Robert D. Van Valin, Jr.
Heinrich-Heine-Universität Düsseldorf & University at Buffalo

1 Aims and motivations

It has been observed that a multitude of the world’s languages can do without explicit formal marking of the concepts of definiteness and specificity through articles (e.g., Russian, Tagalog, Japanese), while other languages (e.g., Lakhota) have very elaborate systems with more fine-grained distinctions in the domains of definiteness and specificity-marking. The main questions that motivate this volume are: (1) How do languages with and without an article system go about helping the hearer to recognize whether a given noun phrase should be interpreted as definite, specific or non-specific? (2) Is there clear-cut semantic definiteness without articles or do we find systematic ambiguity regarding the interpretation of bare noun phrases? (3) If there is ambiguity, can we still posit one reading as the default? (4) What exactly do articles in languages encode that are not analyzed as straightforwardly coding (in)definiteness? (5) Do we find linguistic tools in these languages that are similar to those found in languages without articles?

The papers in this volume address these main questions from the point of view of typologically diverse languages. Indo-European is well represented by Russian, Persian, Danish and Swedish, with diachronic phenomena investigated
Kata Balogh, Anja Latrouite & Robert D. Van Valin, Jr.

in relation to the last two of these. In terms of article systems, they range from Russian, which has no articles, the typical situation in most Slavic languages, to Persian, which has an indefinite article but no definite article, to the more complete systems found in Romance and Germanic languages. The three non-Indo-European languages investigated in this volume, namely Mopan (Mayan), Vietnamese and Siwi (Berber), are typologically quite diverse: Mopan is verb-initial and thoroughly head-marking, Vietnamese is verb-medial and radically isolating, i.e., lacking inflectional and derivational morphology, and Siwi is verb-initial with the signature Afro-Asiatic trilateral roots which are the input to derivational and inflectional processes. What they have in common is the absence of articles signaling (in)definiteness.

2 Article systems and related notions

Chesterman (1991: p.4) points out that “it is via the articles that definiteness is quintessentially realized, and it is in analyses of the articles that the descriptive problems are most clearly manifested. Moreover, it is largely on the basis of the evidence of articles in article-languages that definiteness has been proposed at all as a category in other languages.”

Here, we view definiteness as a denotational, discourse-cognitive category, roughly understood as identifiability of the referent to the speaker, instead of a grammatical (or grammaticalized) category, and therefore we can investigate the means that languages use for indicating definiteness or referential anchoring in general. Natural languages have various means to signal definiteness and/or specificity. Languages differ in their article systems as well as in the functions the set of articles they exhibit may serve. Simple article languages (e.g., English, Hungarian) generally distinguish definite and indefinite noun phrases by different articles, but they may also use their article inventory to code categories other than definiteness (e.g., Mopan Maya). Complex article languages like Lakhota, which exhibits an elaborate and sophisticated system, always mark more than simple (in)definiteness. A great number of languages (e.g., Russian, Tagalog, Japanese) have no or no clear-cut article systems and rely on other means to encode definiteness distinctions.

Most of the languages investigated in this volume belong to the last type. The means they use to help indicate how the referent of a noun phrase is anchored and intended to be interpreted include classifier systems (e.g., Vietnamese, Chuj), clitics (e.g., Romanian), designated morphemes on nouns (e.g., Moksha, Persian) and syntactic position (e.g., Chinese). In certain languages, alongside article sys-
tems and morphosyntactic means, prosody plays a crucial role for the coding of (in)definiteness, for example accent placement in Siwi or tone in Bambara.

2.1 Basic notions: definiteness and specificity

In the cross-linguistic investigation and analysis of article systems and noun phrases, various different but related notions play a key role. In the analysis of various types of definite and indefinite noun phrases, the two most important notions are definiteness and specificity, together with further distinguishing notions of uniqueness, familiarity, discourse prominence and so on. In the following we give a brief introduction to these notions. Our aim is not to provide a detailed discussion of all notions and all theories, but to present an overview of the most important classical analyses relevant to the papers and their main issues in this volume.

2.1.1 Definiteness

The notion of definiteness itself is a matter of controversy, given the different uses of definite noun phrases for anaphoric linkage, relational dependencies, situational/deictic salience or inherently uniquely referring nouns. The notion is used in a variety of ways by different authors. The classical analyses of definiteness distinguish two main lines of characterization: (1) the uniqueness analysis, following works by Russell (1905) and Strawson (1950), and (2) the familiarity account, after Christophersen (1939), Kamp (1981) and Heim (1982).

In Russell’s (1905) analysis, indefinites have existential quantificational force, while definite descriptions\(^1\) are considered referential. Definites assert existence and uniqueness, as illustrated in the logical translation of sentences like (1).

\[(1)\]  
\[\exists x (N(x) \land \forall y (N(y) \rightarrow x = y) \land P(x))\]

- a. there is an N (existence)
- b. at most one thing is N (uniqueness)
- c. something that is N is P

The meaning contribution of the definite article is to signal the existence of a unique referent (a-b), while the head noun provides sortal information of the referent (c). In the Russelian tradition, indefinites are distinguished from definites\(^1\)

\(^1\)These mostly refer to noun phrases with a definite article, e.g., the dog, but other expressions like possessive noun phrases and pronouns are also considered definite descriptions.
in terms of uniqueness, as the predicate (sortal information) applies to exactly one referent. Russell’s highly influential approach has inspired many theories on definiteness; similarly, various approaches point out critical issues in Russell’s theory. The most intriguing issues discussed in the literature are the problem of presuppositionality, the problem of incomplete descriptions and the problem of referentiality. To solve these crucial issues a great number of theories have been proposed over the decades. In Strawson’s (1950) account, existence and uniqueness are presupposed rather than asserted. He claims that if the presupposition fails, the sentence does not bear a truthvalue, i. e., it is neither true nor false. The incompleteness problem, where the definite description does not have a unique referent, inspired several authors (e.g., Strawson 1950; McCawley 1979; Lewis 1979; Neale 1990) to offer various solutions, like contextual restriction and the prominence/saliency approach. The latter was proposed by McCawley (1979) and Lewis (1979), who argues that definite descriptions refer to the most prominent or most salient referent of a given context. Donnellan (1966) argues that definite descriptions have two different uses: an attributive and a referential use. The former use can be characterized similarly to Russell’s account, while the latter use requires a different analysis. Donnellan’s famous example is (2), which can be used in different ways in different situations.

(2) Smith’s murderer is insane.  

(Donnellan 1966: p.285)

In a situation where the murderer is unknown (e. g., at the scene of the crime), the noun phrase ‘Smith’s murderer’ is understood attributively as meaning that whoever murdered Smith is insane. On the other hand, in a different situation where the murderer is known (e. g., at the trial), the noun phrase can be replaced by, for example, he, as it is used referentially, referring to the individual who is the murderer.

The other highly influential classical account of definites represents a different view. These theories follow the work by Christophersen (1939), who accounts for the interpretation of definites in terms of familiarity rather than uniqueness. In his theory, definite descriptions must be discourse-old, already introduced in the given discourse context, and as such known to the hearer. Christophersen’s familiarity account inspired famous theories in formal semantics: File Change Semantics [FCS] of Heim (1982) and the similar Discourse Representation Theory [DRT], which was developed in parallel and introduced by Kamp (1981) and Kamp & Reyle (1993). One of the major contributions of these approaches is the solution for the so-called ‘donkey sentences’ (3a), and further issues of the interpretation of discourse anaphora (3b).
a. If a farmer\(_i\) owns a donkey\(_j\), he\(_i\) beats it\(_j\).

b. A student\(_i\) came in. She\(_i\) smiled.

In both sentences, the indefinite noun phrases can be referred to by an anaphoric expression in the subsequent sentence. Based on such examples, they propose a division of labour between indefinite and definite noun phrases. Indefinites like *a student* introduce new discourse referents, while definite noun phrases like *the student* pick up a referent that has already been introduced, similarly to anaphoric pronouns.

In his 1985 paper, Löbner argues for a relational approach and against the uniqueness approach, claiming that it is rather non-ambiguity that is essential for definiteness. Löbner (1985) distinguishes noun phrases by their type of use. The main distinction is into *sortal* and *non-sortal* nouns, where the latter is further divided into *relational* and *functional* nouns and concepts. Relational nouns include kinship terms (e.g., *sister*), social relations (e.g., *friend*) and parts (e.g., *eye*), while functional nouns are roles (e.g., *wife, president*), unique parts (e.g., *head, roof*), conceptual dimensions (e.g., *height, age*) and singleton events (e.g., *birth, end*). In the analysis of definite descriptions, Löbner (1985) distinguishes semantic and pragmatic definites. For semantic definites “the referent of the definite is established independently of the immediate situation or context of the utterance” (Löbner 1985: p.298), while pragmatic definites are “dependent on special situations and context for the non-ambiguity of a referent” (Löbner 1985: p.298). One- and two place functional concepts (4), as well as configurational uses (5), are considered semantic definites. Löbner claims that statements like (5) are impossible with sortal nouns.

(4) 1-place functional concepts: *the time, the last party, the other girl*, etc.
    2-place functional concepts: *my wife, the author, the president of France*, etc.

(5) He was the son of a poor farmer.  \(\text{(Löbner 1985: ex.17)}\)

As Löbner argues, this distinction is significant in various ways; for example, functional nouns can only take the definite article (with the exception of existential contexts). Further examples he gives are of German cliticization (6), where the cliticized article encodes a semantic definite as opposed to a non-cliticized one. In various languages, there are different articles, often referred to as weak and strong (Schwarz 2019), encoding semantic and pragmatic definites. This distinction can be found, for example, in the Fering (Föhr) dialect of Frisian (e.g., Ebert 1971) and in the Rheinland dialect of German (e.g., Hartmann 1982).
(6) **German:**

   a. Er muß *ins/*in das Krankenhaus.  
      he must *in=the/*in the hospital 
      ‘He must go to hospital again.’
      (from Löbner 1985: ex.52, our glosses)

   b. Er muß *wieder *ins/in das Krankenhaus zurück, aus dem er schon entlassen war.  
      he must again *in=the/in the hospital back from the.DAT he already discharged was 
      ‘He has to go back to the hospital from which he had already been discharged.’
      (from Löbner 1985: ex.54, our glosses)

(7) **Fering (Föhr):**  

   (Ebert 1971: p.161)

   a. Ik skal *deel tu a/*di *kuupman.  
      I must down to theW/theS grocer 
      ‘I have to go down to the grocer.’

      Oki has a horse bought theW/theS horse limps 
      ‘Oki has bought a horse. The horse limps.’

As for the meaning contribution of the definite article, based on the different noun/concept types and their uses, Löbner (1985) argues that the definite article indicates that the given noun must be taken as a functional concept.

### 2.1.2 Specificity

The notion of specificity (see, e.g., von Heusinger 2011) is also defined and characterized in different ways and in relation to a variety of factors. Specificity is generally used to distinguish various readings of indefinites. A generally accepted view is that sentences like (8) can be interpreted in two ways, depending on whether the speaker has a particular entity in mind, referred to by the indefinite noun phrase.

(8) **Mia kissed a student yesterday.**

   1. whoever Mia kissed is a student (non-specific)
   2. there is a specific student whom Mia kissed (specific)

As a linguistic notion, the opposition between the non-specific and the specific readings of indefinites is characterized in relation to a variety of factors. Farkas
(1994) distinguishes referential, scopal and epistemic specificity. Specific indefin- 
ites refer to an individual, and hence can be anaphorically referred back to. With 
respect to the second reading of (8), the sentence could be followed by *He is tall*, 
while this is not possible after the first reading. In relation to other operators, spe-
cific indefinites take a wide scope. The epistemic opposition is very close to (if 
not the same as) the referential opposition, as it is characterized by the fact that, 
by using specific indefinites, the speaker has a referential intention, i.e., they 
have a certain individual in mind (Karttunen 1968; Farkas 1994). In addition to 
Farkas’s (1994) three-way distinction, von Heusinger (2011; 2019) proposes four 
more oppositions, namely partitivity, noteworthiness, topicality and discourse 
prominence. As Enç (1991) argues, specific indefinites are discourse-linked and 
inferable: they refer to a part of a set previously introduced to the discourse. As 
a motivation, she shows that this distinction is overtly marked in Turkish: ac-
cusative marked direct objects are interpreted specifically (9a), while unmarked 
objects are taken as non-specific (9b).

(9)  

   ‘Ali wants to rent a certain piano.’   (Enç 1991: ex.12)  
   (1)  

b. Ali bir piyano kiralamak istiyor.   
   ‘Ali wants to rent a (non-specific) piano.’   (Enç 1991: ex.13)  

The relevance of noteworthiness is often illustrated by the use of the marked 
indefinite *this N* construction. Such examples can only be followed by newswor-
thy/interesting/particular information regarding the noun phrase.

(10) He put this 31-cent stamp on the envelope,   (after Maclaran 1982)  
    a. and only realized later that it was worth a fortune.  
    b. #so he must want it to go airmail.  

Topicality and discourse prominence are also closely related to specificity. In-
definite noun phrases that are topical receive a specific interpretation. This can 
be shown by Hungarian examples, where topicality is syntactically marked by 
placement to a left-peripheral position within the clause (11).

(11) Egy diák be-kopogott az igazgató-hoz.   
    a. student VPRT-knocked the director-ALL  
    ‘A (particular) student knocked at the director’s office.’
The left-peripheral topic position can only host referential and specific noun phrases (e.g., É. Kiss 2002), and hence the indefinite noun phrase can only be in the topic position when it is interpreted specifically.

3 Contributions

The papers in this volume address to different degrees the general questions introduced in §1. Most contributions report on research on different corpora and elicited data or present the outcome of various experimental studies. One paper presents a diachronic study of the emergence of article systems. As mentioned before, the volume covers typologically diverse languages: Vietnamese, Siwi (Berber), Russian, Mopan (Mayan), Persian, Danish and Swedish.

3.1 Languages with articles

If a language is analyzed as having an article, the standard expectation is that it will express either definiteness or indefiniteness. However, the number of papers introducing article-languages in which the determiners do not encode different degrees of identifiability and uniqueness is on the rise (e.g., Lyon 2015). The crucial question is what features an element is required to exhibit to be counted as an article. If the answer is given in line with Himmelmann (2001) and others, then no functional element that does not convey some degree of specificity is counted as an article. If Dryer’s (2014) characterization of articles is adopted, then all functional elements that occur with high frequency in noun phrases, indicate argumenthood and vary for grammatical features are included in the category.

Eve Danziger and Ellen Contini-Morava adopt Dryer’s (2014) view in their contribution Referential anchoring without a definite article: The case of Mopan (Mayan) and investigate all the means that Mopan utilizes in order to evoke relative identifiability and uniqueness. While, based on formal and distributional criteria, the Yucatecan language Mopan exhibits a determiner of the type usually classified as an article, they find that this article does not encode any of the semantic notions of definiteness, specificity and uniqueness. It merely serves to express that a given lexeme is used as an argument in the sentence. In their analysis and explanation, they build upon Dryer’s (2014) definiteness hierarchy and demonstrate that the article itself, as well as the bare nominal form, can occur in any position in Dryer’s definiteness hierarchy. This observation leads to an investigation of exactly what the discourse-pragmatic function of the article is and how it can be calculated. The authors’ conclusion is that the contribution
Nominal anchoring: Introduction

of the article is best characterized by factors such as discourse salience, which contexts or world knowledge may lend even to non-specific indefinites.

In their paper, The specificity marker -e with indefinite noun phrases in Modern Colloquial Persian, Klaus von Heusinger and Roya Sadeghpoor focus on the specificity marker -e and its compatibility with two indefinite markers and investigate the different kinds of indefinite readings that arise. In their experimental pilot studies, they test and provide some support for the hypothesis that the difference in interpretation between the combinations lies in the anchoring of the referents, i.e., in whether the referent is construable as speaker-specific or non-speaker-specific. The studies thereby provide additional evidence for the need to assume a fine-grained approach in the investigation of specificity and referential anchoring (von Heusinger 2002). However, they also show that specificity-unrelated semantic properties like animacy need to be taken into account in the explanation of their results.

The contribution Indirect anaphora from a diachronic perspective: The case of Danish and Swedish by Dominika Skrzypek is the only diachronic study in this volume. The author investigates different kinds of indirect anaphora (associative anaphora, bridging anaphora) as one of the steps in the grammaticalization process towards a definite article from the beginning of the 13th century until the middle of the 16th century. The paper is particularly concerned with the distribution and use of indirect anaphora and the features that the relationship between indirect anaphora and their anchor is based on. Looking at inalienable and other types of indirect anaphora, the author shows that indirect anaphora form a heterogeneous concept and are not easily positioned in the strong-weak definiteness dichotomy. The evidence points to the fact that the definite article did not spread uniformly through indirect anaphora in Danish and Swedish.

3.2 Languages without articles

In article-less languages, the encoding of definiteness is often a complex matter, where various linguistic factors play a role. Japanese and Chinese are both languages that are well known for lacking an article system. In Japanese, argument phrases are marked by case markers (nominative: ga, accusative: wo, dative: ni) or non-case markers like the topic marker wa or the additive marker mo ‘also’. Consequently, definiteness is not straightforwardly grammaticalized, but rather considered an interpretational category (e.g., Tawa 1993), for which classifiers play a crucial role. The same holds for Chinese, which lacks case markers, but exhibits even more numeral classifiers than Japanese. These have been argued by Cheng & Sybesma (1999) and others to play a crucial role for the definiteness
reading of noun phrases, whenever numeral information is missing. However, Peng (2004: p.1129) notes that for indeterminate expressions “there is a strong but seldom absolute correlation between the interpretation of identifiability or nonidentifiability and their occurrence in different positions in a sentence”. Simpson et al. (2011), who study bare classifier definites in Vietnamese, Hmong and Bangla, also find that classifiers are relevant to nominal anchoring. However, the fact that bare noun phrases also seem to be able to receive definite interpretations weakens the claim that classifiers are the morphosyntactic key to definiteness interpretation, and rather points to the fact that a multilevel approach proposed by Heine (1998) is better in explaining how definiteness or specificity interpretations arise.

Walter Bisang and Kim Ngoc Quang, in their study (In)definiteness and Vietnamese classifiers, contribute to our understanding of the classifier language Vietnamese. They investigate which linguistic factors influence the interpretation of phrases with numeral classifiers [CL] in bare classifier constructions as either definite or indefinite and point out the licensing contexts for the different uses and readings of nominal classifiers. They find a striking clustering of definite interpretations with animacy and subject status, whereby definiteness is understood as identifiability in discourse. Indefinite interpretations, on the other hand, are predominantly witnessed in certain sentence types (existential sentences and thetic sentences) and with certain types of verbs (verbs of appearance). A crucial finding is that noun class type, following Löbner (1985; 2011), and factors like animacy and grammatical relation are less important than information structure for the appearance of classifiers in definite and indefinite contexts. Classifiers are shown to be associated with pragmatic definiteness, rather than semantic definiteness, i.e., identifiability rather than uniqueness. Furthermore, the authors provide evidence that contrastive topics, contrastive focus and focus particles correlate with the use of classifier constructions. Similar to the constructions discussed for Persian and Mopan (Mayan) in this volume, the classifier construction in Vietnamese can be once more viewed as a construction whose final interpretation depends, on the one hand, on discourse prominence and, on the other hand, on features of the morphosyntax-semantics interface that are well known for contributing to the overall saliency of a phrase.

In her contribution, Accent on nouns and its reference coding in Siwi Berber, Valentina Schiattarella investigates definiteness marking in Siwi Berber, an indigenous Berber language spoken in Egypt. In Siwi, a language without articles, it is claimed that the placement of accent on the last syllable versus the penultimate syllable encodes indefiniteness and definiteness respectively, i.e., the ac-
cent on the last syllable is generally assumed to encode indefiniteness and the accent on the penultimate syllable to encode definiteness. This default interpretation can be overridden, as Schiattarella shows in her paper. She analyzes various corpus data from spontaneous discourse and guided elicitations to further examine the role of various morphosyntactic means (e.g., possessive constructions, demonstratives, prepositions and adpositional phrases) as well as pragmatic aspects (e.g., anaphoricity, familiarity, uniqueness, reactivation and information structural considerations) in influencing the interpretation of noun phrases. The author, furthermore, finds that right- and left-detached constructions or the appearance of a demonstrative, a possessive marker or relative clause in postnominal position influences the interpretation.

Olga Borik, Joan Borràs-Comes and Daria Seres, in *Preverbal (in)definites in Russian: An experimental study*, present an experimental study on Russian bare nominal subjects, and investigate the relationship between definiteness, linear order and discourse linking. Given that Russian lacks articles and has very flexible word order, it is widely assumed that (in)definiteness correlates with the position of a noun phrase in the clause, i.e., preverbal position is associated with a definite reading and postverbal with an indefinite interpretation. The authors experimentally verify that this correlation basically holds, but they also find that speakers accept a surprising number of cases in which a preverbal NP is interpreted as indefinite, which leads to the conclusion that Russian bare nouns are basically indefinite. The unexpected correlations between position and interpretation lead to further investigations of the relevant factors involved and the suggestion that, regardless of topicality, discourse linking principles following Pesetsky (1987) and Dyakonova (2009) facilitate the use of indefinite nominals in the unexpected preverbal position.

### 3.3 Summary

The papers in this volume deal with pragmatic notions of definiteness and specificity. The studies presented here provide the following findings regarding our initial motivating questions. On the issue of how languages with and without articles guide the hearer to the conclusion that a given noun phrase should be interpreted as definite, specific or non-specific, the studies in this paper argue for similar strategies. The languages investigated in this volume use constructions and linguistic tools that receive a final interpretation based on discourse prominence considerations and various aspects of the syntax-semantics interface. In case of ambiguity between these readings, the default interpretation is given by factors (e.g., familiarity, uniqueness) that are known to contribute to the salience
of phrases, but may be overridden by discourse prominence. Articles that do not straightforwardly mark (in)definiteness encode different kinds of specificity. In the languages studied in this volume, whether they have an article system or not, similar factors and linguistic tools are involved in the calculation process of interpretations.

Acknowledgments

The volume contains revised selected papers from the workshop entitled *Specificity, definiteness and article systems across languages* held at the 40th Annual Conference of the German Linguistic Society (DGfS), 7-9 March, 2018 at the University of Stuttgart. We very much appreciate the contributions of all participants in the workshop, who enriched the event with their presence, questions, presentations and comments. Special thanks to the DFG for all financial support and their funding of our project D04 within the SFB 991 in Düsseldorf, which made the workshop and this editing work possible. Many thanks to the series editors Philippa Cook, Anke Holler and Cathrine Fabricius-Hansen for carefully guiding us through from the submission to the completion of this volume.

References


1 Nominal anchoring: Introduction


