

Handbücher zur Sprach- und Kommunikations- wissenschaft

Handbooks of Linguistics
and Communication Science

Manuels de linguistique et
des sciences de communication

Mitbegründet von
Gerold Ungeheuer

Herausgegeben von / Edited by / Edités par
Armin Burkhardt
Hugo Steger
Herbert Ernst Wiegand

Band 20.1

Walter de Gruyter · Berlin · New York
2001

ulary. (Cushitic Language Studies 2). Hamburg: Buske.

Hutchisson, Don. 1986. "Sursurunga pronouns and the special uses of quadral number". In: Ursula Wiesemann (ed.) *Pronominal systems* (Continuum 5), Tübingen: Narr. 217–55.

Kibrik, Alexandr E. 1991. "Organising principles for nominal paradigms in Daghestanian languages: comparative and typological observations". In: Frans Plank (ed.) *Paradigms: the economy of inflection* (Empirical Approaches to Language Typology 9), Berlin: Mouton de Gruyter, 255–74.

Laidig, Wyn D. & Laidig, Carol J. 1990. "Larike pronouns: duals and trials in a Central Moluccan language". *Oceanic Linguistics* 29: 87–109.

Lichtenberk, Frantisek. 1983. *A grammar of Manam* (Oceanic Linguistics Special Publication 18). Honolulu: University of Hawaii Press.

Masica, Colin P. 1991. *The Indo-Aryan languages*. Cambridge: Cambridge University Press.

Mithun, Marianne. 1988. "Lexical categories and the evolution of number marking". In: Michael Hammond & Michael Noonan (eds.) *Theoretical morphology: approaches in modern linguistics*. San Diego: Academic Press, 211–34.

Newman, Paul. 1990. *Nominal and verbal plurality in Chadic* (Publications in African Languages and Linguistics 12). Dordrecht: Foris.

Plank, Frans. 1989. "On Humboldt on the dual". In: Roberta Corrigan, Fred Eckman & Michael

Noonan (eds.) *Linguistic categorization* (Amsterdam Studies in the Theory and History of Linguistic Science 61). Amsterdam: Benjamins, 293–333.

Priestly, T. M. S. 1993. "Slovene". In: Bernard Comrie & Greville G. Corbett (eds.), *The Slavonic languages*. London: Routledge, 388–451.

Reid, Nicholas J. 1990. *Ngan'gityemmerri: A Language of the Daly River region, Northern Territory of Australia*. Unpublished PhD thesis, Australian National University, Canberra.

Saussure, Ferdinand de. 1971. *Cours de linguistique générale* (publié par Charles Bally et Albert Sechehaye avec la collaboration de Albert Riedlinger). Paris: Payot. [Third edition, first edition 1916.]

Smith-Stark, T. Cedric. 1974. "The plurality split". *Chicago Linguistic Society* 10: 657–71.

Steever, Sanford B. 1987. "The roots of the plural action verb in the Dravidian languages". *Journal of the American Oriental Society* 107: 581–604.

Stone, Gerald. 1993. "Sorbian (Upper and Lower)". In: Bernard Comrie & Greville G. Corbett (eds.), *The Slavonic languages*. London: Routledge, 593–685.

Tiersma, Peter M. 1982. "Local and general markedness". *Language* 58: 832–49.

Watkins, Laurel J. 1984. *A grammar of Kiowa*. Lincoln and London: University of Nebraska Press.

Greville G. Corbett, *University of Surrey*
(United Kingdom)

62. Articles

1. Two approaches to the typology of articles
2. Identifying articles
3. Articles derived from demonstratives
4. Articles from other sources
5. Co-occurrence with other determiners
6. Special abbreviations
7. References

1. Two approaches to the typology of articles

The typology of articles has been approached in essentially two ways. The approach adopted by Grasserie (1896) and Krámský (1972) begins with a definition of definiteness and then identifies and classifies various grammatical phenomena which express definiteness in natural languages. The resulting typology is not limited to articles but includes other grammatical phenomena such as word order and verbal agreement. Semantically,

such a typology is based on the simple bilateral opposition 'definite' vs. 'indefinite', which is assumed to be universal. Consequently, the typological parameters are exclusively formal. Krámský (1972), for example, subdivides his list of definiteness expressing phenomena according to the following two parameters: the morphological shape of the definiteness markers (independent word, clitic, affix, non-segmental phenomena (order, stress)) and the number of definiteness markers (definite and indefinite, only definite, only indefinite).

The second approach, pioneered by Greenberg (1978), is based on the observation that articles are historically derived from other elements such as demonstratives and numerals. That is, articles are viewed as stages in the adnominal grammaticisation of these elements, the typology of articles then

being concerned with identifying different paths of grammaticisation in which articles occur and classifying the different stages of their development on the basis of semantic and formal criteria.

Probably the best-known grammaticisation path for articles is Greenberg's somewhat rudimentary grammaticisation path for articles derived from demonstratives:

- (1) DEM → DEF.ART → SPEC.ART → NOUN
MARKER

This path may be used briefly to preview some major features of the grammaticisation approach to the typology of articles and to highlight some salient differences with regard to the definiteness-based approach.

Most importantly, from a grammaticisation point of view, definiteness, though undeniably of central importance to the grammar and typology of articles, is only one of a number of meanings that can be conveyed by articles. Thus, for example, specific articles indicate specificity rather than definiteness. Definiteness and specificity are only two salient (and crosslinguistically well attested) stages on the grammaticisation path in (1). Given a sufficiently fine-grained set of criteria, many more stages could be distinguished. Thus, on the one hand, the grammaticisation approach provides a more detailed typological grid for articles and includes a variety of phenomena not considered in the definiteness-based approach (further details in § 3 and 4). On the other hand, the grammaticisation approach excludes some of the phenomena dealt with in the definiteness-based approach, in particular those grammatical structures which appear to be sensitive to definiteness but are not part of a grammaticisation path for articles (such as word order, case markers or verbal agreement).

Furthermore, the grammaticisation approach highlights the problem of distinguishing demonstratives from articles (and various types of articles from each other). In order to establish a grammaticisation path for articles, explicit and operational criteria have to be provided for determining that a given element is no longer a demonstrative but a definite article. This problem is addressed in § 2.

It should be noted that both approaches converge in that they both address the issues of classifying definite and indefinite articles according to their different morphosyntactic shapes (see in particular § 3.2) and according

to the paradigmatic relations they enter into (see §§ 4.1 and 5).

The present overview of the typology of articles basically follows the grammaticisation approach as sketched in Greenberg (1978) and further elaborated in Himmelmann (1997) and Laury (1997). It is primarily concerned with articles derived from demonstratives, the most common and widespread type in natural languages.

2. Identifying articles

This section presents some formal (morphosyntactic) and semantic criteria for distinguishing articles from related elements, in particular the elements they are historically derived from, such as demonstratives and numerals (some of these criteria are also relevant to distinguishing different types of articles).

Two essential, though not sufficient formal criteria are as follows. First, articles are grammatical elements which occur only in nominal expressions. Second, their position within such expressions is fixed (i. e. they occur either to the left or to the right of a noun but not alternatively on either side). At first sight, these may seem to be rather trivial criteria. But note that they distinguish articles from, for example, demonstratives in many languages. Demonstratives (more precisely: deictic elements) may appear in at least four different syntactic functions, i. e. as pronouns, as adverbials, as predicates, and as adnominal modifiers. In the first three of these functions they are not in construction with a noun (cf. Fillmore 1982: 47f., Diessel 1999, Himmelmann 1997: 125–130). Furthermore, in many languages (e. g. Swahili, Latin, many Australian languages) demonstratives may occur on either side of a noun when used adnominally (cf. Dryer 1989). Thus, an element which can occur by itself, or one the position of which in nominal expressions is not fixed is, in general, not an article.

A further useful heuristic for identifying articles is obligatoriness in grammatically definable contexts. Such contexts include superlatives (*the greatest singer*), nouns taking complement clauses (*the fact that they lost the game*), and the fact that count nouns cannot be used in core argument positions without a marker for definiteness or specificity.

High textual frequency, often invoked as an argument for claiming that a particular el-

ement is an article, is an important but equivocal indicator of articlehood. Unequivocal articles (when not bound formatives) are without doubt high frequency items, found generally at the very top of frequency statistics for a given language sample. Thus, claims that a low frequency item should be regarded as an article are generally met with well justified suspicion. However, the mere fact that a demonstrative or numeral is used more frequently in one language or historical period than in another does not necessarily show that the element in question has become grammaticised as an article. In order for this to be the case it has to be shown that the increase in overall textual frequency correlates with a substantial difference in the semantics and pragmatics of the grammaticising element (cf. Himmelmann 1996: 210–218 and 1997: 28–33, 192–194 for further details on this methodological issue and problems regarding arguments based solely on frequency).

Formal criteria alone are not sufficient for providing a useful and operational definition of articles since there are other kinds of grammatical elements in nominal expressions which share essentially the same formal characteristics (case markers or classifiers, for example). They have to be complemented by semantic criteria. These differ strongly according to the type of article and its historical source. In the remainder of this section only the semantic criteria for distinguishing demonstratives from articles are discussed. The much less well known criteria relevant to other types of articles are briefly mentioned in § 4.

Put simply, articles are distinguished from demonstratives by the fact that they can be used in semantic and pragmatic contexts in which demonstratives cannot be used. The discussion here follows Hawkins (1978, 1991) who, building on the classic work of Christophersen (1939), has presented a very useful systematisation of article use in English which in turn provides a crosslinguistically applicable grid for identifying definite and specific articles semantically.

The first step in establishing such a grid is to identify those contexts in which demonstratives can be used in all languages. Himmelmann (1996) argues that there are four such contexts. The following three of these (illustrated from English Pear Stories, cf. Chafe 1980) are well-known:

- (2) **Situational use**
This guy behind you waits to get back to his seat (referring to a person present in the utterance situation)
- (3) **Discourse-deictic use**
... and that's the end of that story (referring to a preceding stretch of discourse)
- (4) **Tracking use** (also called anaphoric use)
... and a man comes along with a goat, and this goat obviously is interested in the pears

The fourth universal demonstrative use, **recognitional use**, is less widely known. This use is characterised by the fact that the intended referent has to be identified via specific, but presumably shared, knowledge. It can always be (and in fact often is) accompanied by a *you know? or remember?*-type of tag question.

- (5) **Recognitional use**
 - (a) *those dusty kind of hills that they have out here by Stockton and all*
 - (b) *hitting one of those bounce-back things, you know, the little thing that had elastic, and it has a ball*

Note that in the preceding examples the distal demonstrative *those* occurs in first mentions. Neither the hills nor the bounce-back things have been mentioned before. They are also not visibly part of the actual utterance situation.

Articles derived from demonstratives may or may not be useable in these four contexts. The crucial distinguishing feature, however, is that they are consistently used in some additional contexts in which demonstratives must not be used. For **definite articles** two contexts are of particular importance. One is **larger situation use**, the first mention of entities that are considered to be unique, and hence generally identifiable, in a given speech community. This use is characterised by the fact that the intended referent has to be identified via general knowledge (e. g., *the sun, the Queen, the pub*). The other is **associative-anaphoric use**, i. e. the first mention of an entity that is not unique per se but with respect to a previously mentioned referent, as in:

- (6) *The man drove past our house in a car. The exhaust fumes were terrible.*

As stated above, demonstratives cannot be used in these two contexts. In (6) it is impos-

sible to replace *the* in *the exhaust fumes* by either *these* or *those*. Similarly, one cannot say *Her face was burnt by this sun* if *sun* is mentioned for the first time in a given conversation.

Specific articles (which are very common in Niger-Congo and Austronesian languages, among others) may occur in all of the preceding contexts. In addition, however, they may be used in some contexts in which neither demonstratives nor definite articles may occur. A useful diagnostic context in this regard is the introduction of a new participant into the universe of discourse as in the following Tagalog example (so-called **specific-indefinite use**):

- (7) Tagalog (Bloomfield 1917: 32/31)
doón ay na-kita nilá
 DIST.LOC PM REAL.STAT-see 3PL.POSS
ang isá-ng ma-lakí-ng higante
 SPEC one-LK STAT-big-LK giant
 ‘There they saw a great giant ...’

Put more generally, the crucial difference between specific articles on the one hand and demonstratives and definite articles on the other is that in the case of specific article use it is not necessary that the speaker assumes that the intended referent is identifiable for the hearer.

3. Articles derived from demonstratives

There are two basic types of articles derived from demonstratives. One type may be called **phrasal article**. These articles are found at the (left or right) periphery of a nominal expression. In general, they occur only once per nominal expression. The other type may be called **linking article**. These articles are found in between the constituents of a complex nominal expression. They may occur several times in a single nominal expression.

Phrasal articles and linking articles develop from different syntactic constellations. The former derive from a syntagm consisting of DEM-(X)-N or N-(X)-DEM (X representing optional adnominal modifiers such as adjectives, quantifiers, etc.), while the latter derive from a syntagm of the shape N-DEM-X or X-DEM-N. Thus, one major difference between phrasal and linking articles pertains to the fact that linking articles only occur in complex nominal expressions, while phrasal articles occur in both simple and complex nominal expressions.

3.1. Linking articles

Linking articles (which are often also called linkers, ligatures, or connectors) may be indeclinable particles which occur between the major constituents of a nominal expression, as in the following example from Tagalog:

- (8) Tagalog (Bloomfield 1917: 24/37)
apat na malalim na balón
 four LK deep LK well
 ‘four deep wells’

Or they may agree with the noun, as in the following Albanian examples:

- (9) Albanian (Buchholz & Fiedler 1987: 326)
shok-u i mirë
 friend-DEF.NOM.SG.M LK.NOM.SG.M good
 ‘the good (male) friend’
- (10) Albanian (Buchholz & Fiedler 1987: 326)
shoqja e mirë
 friend:DEF.NOM.SG.F LK.NOM.SG.F good
 ‘the good (female) friend’

The preceding examples illustrate the use of linking articles with quantifiers and adjectives. Other typical contexts of use include relative clauses and (much more rarely) demonstratives:

- (11) Tagalog
yuing mama na bàbaril sa
 iyon-ng mama’ na RED1-baril sa
 DIST-LK man LK RED1-gun LOC
kanyá
 kaniyá
 3SG.DAT
 ‘that man who was going to shoot at him’

In both Albanian and Tagalog the linking articles are strongly grammaticised. Their use depends exclusively on grammatical factors such as the kind of adnominal modifier. Semantic and pragmatic factors such as definiteness or specificity do not play a role; cf. example (7) above and:

- (12) Albanian (Buchholz & Fiedler 1987: 326)
(një) shok i mirë
 one friend:INDEF.M LK.NOM.SG.M good
 ‘a good (male) friend’

In Albanian, the linking article is also used with genitive modifiers:

- (13) Albanian (Buchholz & Fiedler 1987: 356)
 roman-i i tretë
 novel-DEF.NOM.SG.M LK.NOM.SG.M third
 i autor-it
 LK.NOM.SG.M author-DEF.GEN.SG.M
 ‘the author’s third novel’

In Tagalog the linking article does not form a constituent with either of the two elements it ‘links’. In Albanian, however, there is good evidence that it forms a constituent with the modifier, either as a proclitic (in the case of genitive modifiers) or as a prefix (in the case of adjectives). Thus, adjectives, for example, also occur with a linking article in predicative function, as in:

- (14) Albanian (Buchholz & Fiedler 1987: 199)
 vajz-a është
 girl-DEF.NOM.SG.F be.3SG.PRS
 e bukur
 LK.NOM.SG.F pretty
 ‘the girl is pretty’

The development of linking articles roughly follows the bondedness hierarchy proposed by Foley (1980): they occur first with attributes such as relative clauses and adjectives, then with quantifiers, and finally with demonstratives (cf. Himmelmann 1997: 172–183). Gamillscheg (1937) identifies examples such as Latin *porcus ille silvaticus* ‘pig that feral’ as a possible source construction.

The development of linking articles appears to be totally independent from the development of phrasal articles. In several languages, including Tagalog and Albanian, linking articles freely co-occur with phrasal articles in the same construction.

3.2. Phrasal articles

There are two basic types of phrasal articles. One type, called **NP-article** here, occurs in all types of nominal expressions, including nominal expressions functioning as modifiers within complex nominal expressions (as in *the end of the movie*). The definite articles in European languages are prototypical examples of NP-articles.

The other type, called **complement article** (CMA) here, occurs only once per clause-level nominal expression (i.e. a nominal expression which is in construction with a finite verb or one which forms a major constituent in a non-verbal clause). Compare the following examples from the Papuan language Hua:

- (15) Hua (Haiman 1977: 58)
finaroga-mo vie
 fight-ALL-CMA go.3SG
 ‘He went to the/a fight.’
- (16) Hua (Haiman 1977: 58)
*finaroga (*mo) de(-mo)*
 fight-ALL man(-CMA)
 ‘a man for fights/a fighting man’
- (17) Hua (Haiman 1978: 568)
*bura de-ma? (*mo) fu(-mo)*
 that man-GEN pig(-CMA)
 ‘that man’s pig’

In (15) the nominal expression *finaroga* is the allative complement of the verb ‘go’ and hence marked by the complement article *mo*. In (16) the same nominal expression functions as a complement of the noun *de* ‘man’. In this construction, use of the complement article *mo* is impossible. Similarly in (17), where *bura dema?* ‘that man’s’ functions as a possessive modifier of ‘pig’. Further details on the complement article may be found in § 3.2.2.

There is no well-established terminology for many of the phenomena considered in this section, including the distinction between complement articles and NP-articles. Alternative terms for complement articles include ‘connective particle’ or ‘phrase marker’.

3.2.1. NP-articles

The usage conditions for NP-articles vary across languages with regard to many minor details. For the two best known types of NP-articles, the definite and the specific article, the following contexts have been identified as particularly prone to crosslinguistic variation (cf. Greenberg 1978: 64–66 and Krámský 1972: 74–89): a) use with proper names and vocatives; b) use in adpositional phrases (cf. Himmelmann 1998); c) use with other determining elements such as demonstratives and possessors (see § 5). For specific articles there are the following additional grammatical contexts (cf. Greenberg 1978: 66–69): a) the noun functions as the object of a negated verb; b) the noun functions as a nominal predicate; c) the noun appears as part of a compound.

With regard to formal aspects, NP-articles vary along at least two parameters: a) the degree of fusion between article and noun; b) the number of articles in a given system (cf. Krámský 1972: 74–165 and Bechert 1993).

NP-articles typically are phrasal proclitics, occurring on the leftmost periphery of a

noun phrase (as in *the three little boys*). Less commonly, they appear as phrasal enclitics as in Balinese *karanjang gede ne* ‘the big basket (lit. basket big the)’. Very rarely, they are second position clitics, i. e. they appear after the first constituent of a noun phrase. In Bulgarian and Macedonian, for example, the definite article occurs after the noun if the NP consists of just a noun, but if the noun is preceded by one or more modifiers, it occurs immediately after the first modifier:

- (18) Bulgarian (Scatton 1984: 164–167, 314f.)
kníga=ta
 book:SG.F=DEF.SG.F
 ‘the book’
dvé=te nóvi knígi
 two.F=DEF.PL.F new:PL book:PL.F
 ‘the two new books’

Furthermore, NP-articles can be suffixed to the noun, a phenomenon common in North Germanic languages (e. g. Danish *hus* ‘house’ vs. *hus-et* ‘the house’) and also found in Rumanian (e. g. *lup* ‘wolf(M)’ vs. *lup-ul* ‘the wolf’, *casă* ‘a house (F)’ vs. *casa* ‘the house’) and Albanian. Prefixed articles appear to be rare. The definite articles in Berber and Semitic languages possibly belong into this category.

An extremely rare phenomenon is the consistent enclisis of an article to the preceding constituent. This phenomenon is found in North Wakashan (Boas 1911, 1947 and Anderson 1992: 23–37) and the neighbouring Tsimshian languages (Mulder 1994: 30–49). In the following example from Kwakw’ala (North Wakashan) the NP-articles =*i* and =*a* are both glossed as DIST because they convey deictic meanings, a point to which we will return shortly. Here it is only important to register that these articles are phonologically part of the preceding word while grammatically forming a constituent with the following one.

- (19) Kwakw’ala (Boas 1911: 554)
lá:’lai: dú:x^waʔlaxa
 la:-’la=i du:q^w-aʔla=χ=a
 go-RPRT-DIST see-accomplish-OBJ-DIST
 ‘nəqá:c’aqi: x^wá:k’una
 ’nəq-c’aq=i x^wa:k^w-n=a
 ten-LONG-OBJECT-DIST canoe-??-DIST
məxí:s a:q
 məx-i:s la:q
 hollow_thing-beach go-DIST.VIS.OBJ
 ‘There he discovered ten canoes on the beach.’

Also extremely rare and in need of further investigation are the claims that an NP-article is realised only suprasegmentally, i. e. as a modification of the stress pattern of a noun. Krámský (1972: 186) mentions Ossetic as a language where this phenomenon is found. Another example is the so-called ‘definitive accent’ in Tongan (Churchward 1953: 25–27). This ‘accent’, however, appears to be functionally a demonstrative according to the criteria stated in §2 since, judging from Churchward’s examples, it occurs in typically recognitional and anaphoric contexts (Tongan also has a proclitic specific article *e*). Furthermore, the ‘definitive accent’ formally occupies an otherwise empty slot in the paradigm of enclitic demonstratives (cf. Clark 1974).

As for the number of NP-articles being part of a single system, a variety of phenomena may be distinguished. Typically, there is only one NP-article in a given system which may be either an indeclinable particle (as English *the* or Tagalog *ang*) or it may agree in number, gender and/or case with the noun (as in French, Albanian, Bulgarian, Standard German, etc.). More complex systems arise when NP-articles convey additional information about the referent (**classifying articles**) or speakers’ attitude (**emotional or honorific articles**). Honorific articles occur primarily with proper nouns (they are widespread Western Austronesian, e. g., Tagalog *si*, Balinese *ni*). Emotional articles are found in certain Polynesian languages where they convey sympathy or belittlement (e. g. Samoan *nāi teine* ‘the poor/dear/few girls’; cf. Mosel & Hovdhaugen 1992: 264–267). Krámský (1972: 93–96) reports a fairly elaborate system of classifying articles for Ponca (Siouan).

In a totally different type of system with multiple NP-articles the choice between two (or more) different NP-articles depends on the syntactic construction. For example, in North Germanic languages the definite article for unmodified nouns is a suffix (cf. Braunmüller 1982: 222–242). However, if the noun is preceded by modifiers then a different form of the definite article, i. e. a phrasal proclitic, has to be used (e. g., Danish *mand-en* ‘the man’ and *den gamle mand* ‘the old man’). In some of these languages, both forms of the definite article are used together in complex nominal expressions. Thus, Swedish *den gamle mann-en* ‘the old man’ includes both a proclitic and a suffixed article.

A third type of system with multiple NP-articles arises when articles encode meanings typically associated with demonstratives such as visibility (VIS vs. INVIS) or distance from a deictic center (PROX, MED, DIST). These articles may be called **deictic articles**. They have been reported for Salish and North Wakashan (according to Boas (1947: 260) the two clitics =*i* and =*a* in example (19) above differ with regard to visibility) and several Austronesian languages, for example Tsou (Tung 1964: 218). To date, many aspects of these articles are not very well understood, including the precise nature of the semantic differences between the members of a given paradigm and between these articles and ‘true’ demonstratives.

Here it will suffice to illustrate the phenomenon from one language. In Lushootseed (Salish) the paradigm of determiners includes the following three elements: *ti* ‘the (definite)’, *tiʔəʔ* ‘this’, and *tiʔit* ‘that’ (glosses from Hilbert & Hess 1977: 5), the latter two obviously being derived in some way from the first one (cf. Montler 1986: 224–236 for a componential analysis of Salish determiners). Although the glosses suggest that *tiʔəʔ* and *tiʔit* are demonstratives, their use appears to be similar to specific articles. That is, in narrative texts they are used not only in typically demonstrative contexts but also in contexts which appear to be interpretable as instances of larger situation, associative-anaphoric or even specific-indefinite use (for the latter cf. also Jelinek & Demers 1994: 731 f.). Thus, the following example contains the first mention of the two protagonists of a traditional story, marked by *tiʔəʔ* (for want of a better gloss here simply glossed as THIS):

- (20) Lushootseed (Hilbert & Hess 1977: 14)
tu-huyu-cut-əx^w tiʔəʔ
 PAST-make-self-NOW THIS
[sə]saʔliʔ ləg^w-ləg^wəb
 two (humans) RED-youth
 ‘Two youths prepared themselves.’

Some of the multiple NP-article systems reported in the literature turn out, upon closer inspection, to contain demonstratives. Thus, it has been claimed for a variety of German dialects (Heinrichs 1954: 85–103, Hartmann 1982), North Frisian (Ebert 1970) and some Catalan dialects (references and some discussion in Selig 1992: 191) that two definite articles exist in these languages and dialects. One of these so-called definite articles, however, is used only in anaphoric contexts. Hence by

the criteria used here, it is not an article but a high-frequency demonstrative (cf. Himmelmann 1997: 54–56, 100).

3.2.2. Complement articles

In addition to the defining feature of complement articles stated at the beginning of this section – i.e. that they occur only once per clause-level nominal expression – these articles also tend to occur after subordinate clauses expressing presupposed information. Thus, in Hua a complement article is commonly found after conditional and temporal adverbial clauses:

- (21) Hua (Haiman 1980: 498/30)
zuʔ-roga-mo bau-ma-mo
 house-ALL-CMA be-REL.DES-CMA
kgaiʔ-hiʔ d-geta haviʔ
 2SG-BEN 1SG.POSS-mind think
haviʔ hu-da
 think do-1SG
 ‘When I am at home I just think about you all the time.’

For a more extensive discussion of complement articles in Papuan languages, representing different stages in their grammaticisation, see Reesink (1994). Complement articles are in all likelihood not restricted to Papuan languages. Elements with similar distribution occur in some American Indian languages (Cayuga, Lakota) and possibly also in Haitian Creole (Lefebvre & Massam 1988).

4. Articles from other sources

This section briefly reviews articles derived from sources other than demonstratives. The review is brief not only because the phenomena to be discussed here are less widespread and less frequently attested in natural languages but also because there are, to date, no in-depth studies on which well-supported generalisations could be based.

4.1. Articles derived from ‘one’

Indefinite articles derive quite generally from the numeral ‘one’. In fact, in many languages the numeral ‘one’ and the indefinite article are segmentally identical (e.g. French *un/lune*, written German *ein/eine*, Turkish *bir*). And even if indefinite article and numeral are not segmentally identical, they generally do not co-occur with each other (**a one book* is not well-formed in English; cf. also Moravcsik 1969: 84). An exception is Sinhala where what appears to be an indefinite article freely

co-occurs with numerals, including the numeral for 'one'. Examples: *pota-k* 'a book (book.SG-INDEF)', *pot eka-k* 'one book (book.PL one-INDEF)', *pot tuna-k* 'three books (book.PL three-INDEF)', *mee pot tuna* 'those three books'.

When numeral and article are segmentally identical, the major formal criterion invoked in the literature for distinguishing them is that the article is unstressed. This, however, is a rather weak criterion since in most languages numerals do not have to be stressed. Consequently, in some instances it is controversial whether it is actually possible and useful to distinguish between article and numeral (for German, see Bisle-Müller 1991: 100–116, among others). In other instances stronger formal evidence is available for making such a distinction. In Turkish, for example, numeral and article are claimed to occupy different syntactic positions in complex nominal expressions (Lewis 1967: 54). The article intervenes between modifier and noun (*büyük bir tarla* 'a large field') while the numeral precedes the adjective (*bir büyük tarla* 'one large field').

Another kind of formal evidence for indefinite articles is provided by the rather rare phenomenon that an indefinite article derived from the numeral 'one' may be used in plural expressions. In at least one language, Spanish, special plural forms of such an indefinite article exist (cf. *un libro* 'a book' vs. *unos libros* 'some/a few books'). It should be noted, however, that indefinite articles derived from 'one' are typically restricted to the use with singular count nouns, a major difference between definite and indefinite articles. Exceptions to this generalisation include the indefinite articles in Turkish and, possibly, Sinhala. Furthermore, it is occasionally argued that another quantifier (e.g. *some* in English) functions as a plural indefinite article (Chesterman 1991: 44 f.). Fijian and some Polynesian languages are said to have plural indefinite articles derived from sources other than 'one'.

Givón (1981) is the only crosslinguistic study to date that attempts to identify different stages in the grammaticisation of indefinite articles, thereby also providing semantic criteria for distinguishing numerals from articles. One important criterion is the consistent use of the latter in nonspecific-indefinite contexts such as *I am looking for a book on math, do you have any?* Another one is the use of the indefinite article with predicate nominals as in *He is a language profes-*

sor. Furthermore, indefinite articles may occasionally be used in some kinds of generic statements such as *A dodo likes peanuts* (see Chesterman 1991: 32–38 for references and discussion).

As to the formal realisations of indefinite articles, the crosslinguistic variation is much more restricted than in the case of NP-articles. To date, no systems containing multiple indefinite articles have been identified. The major crosslinguistic variation pertains to the position of the indefinite article (before or after the noun) and to its degree of fusion with the noun (indefinite article are clitics in European languages but suffixes in Persian and Sinhala).

It is common to think of definite and indefinite articles as a 'natural pair', i.e. as occurring together in one morphosyntactic paradigm. Crosslinguistically, however, this is the exception rather than the rule (see Moravcsik 1969: 85–89). There are many languages with definite articles lacking indefinite articles (classical Greek, the Celtic languages, Bulgarian, many modern Arabic dialects, etc.). Conversely, there are a few languages with indefinite articles but no definite articles (e.g., Persian, Sinhala). And even if both a definite and an indefinite article are found in a particular language, they do not necessarily exhibit similar morphosyntactic characteristics. In several languages, for example, one article precedes the noun, the other follows it (e.g., Classical Arabic, North Germanic languages).

4.2. Pronominal articles

In many languages, personal pronouns can be used adnominally as in *we students*. This use, however, is generally severely constrained by grammatical and semantic factors and rather infrequent overall. In a few languages, including most of the languages spoken in central Australia (e.g., Mparntwe Arrernte (Wilkins 1989), Yankunytjatjara (Godard 1985)), third person pronouns are regularly used in anaphoric mentions, e.g.:

- (22) Mparntwe Arrernte (Wilkins 1989: 129)
- | | | |
|--------------------------|-----------|--------------|
| <i>Artwe itne</i> | <i>no</i> | |
| man | 3.PL.NOM | NEG |
| <i>ahel-irre-ke</i> | | <i>artwe</i> |
| angry-INCH-PAST.COMPL | man | |
| <i>mperikerre ikwere</i> | | |
| white | 3.SG.DAT | |
- 'The men didn't become aggressive towards the white man.'

In these languages, then, third person pronouns are used more frequently and in more kinds of contexts than in other languages. This can be interpreted as the beginning of the grammaticisation of an article from a pronoun. Given that third person pronouns are also generally derived from demonstratives, this grammaticisation path for articles is clearly closely related to the one leading directly from demonstratives to articles.

A fully grammaticised pronominal article is found in Nama (Khoisan) where every specific nominal expression is accompanied by a pronominal clitic (in phrase-final position). In the following example both the 1.SG pronoun and the word for 'desert' are followed by such a pronominal clitic, agreeing with the phrasal head in person, number and gender:

- (23) Nama
ti-ta ke k!aro-ŋ!hu-p ŋ!â
 1.SG-1.SG SUBJ desert-land-3.SG.M LOC
nî kl?an-mâ.
 FUT.PFV smoke-stay
 'I will pitch in the desert.'

Since there are clitics for all persons and any noun may, in principle, be followed by any clitic (e. g. *si-?ao-khom* (1.NONSG-man-1.DU.M) 'we two men'), it is sometimes said that in Nama nouns are 'inflected' for person. But note that these elements are clearly clitics (and not affixes) and that their overall distribution corresponds to a strongly grammaticised specific article.

4.3. Possessive articles

There is also evidence that in several languages possessive pronouns can be used well beyond the typical contexts of use for possessive pronouns. This is particularly clear in instances where there is no antecedent for the pronoun, as in the following example from Indonesian:

- (24) Indonesian
karena sungai-nya keruh
 because river -3SG.POSS muddy
 '(We couldn't take a bath) because
 the river was muddy ...'

The river referred to here is the river known to everybody in the speech community. It is clear that there is no possessor for the river (i. e. it would be wrong to translate *sungainya* with 'his/her/its river'). Example (24), then, shows a larger situation use of a possessive pronoun, a context highly characteristic of definite and specific articles.

Similar extended uses of possessive pronouns have been reported for several Uralic

and Turkic language in which both 3.SG and 2.SG possessive pronouns can be used in this way (cf. Krámský 1972: 173 f., Bechert 1993: 31–36). But in these languages as well as in Indonesian it does not appear to be the case that use of a possessive pronoun is obligatory in all (or a clearly defined subset of) definite or specific contexts. Further research is required to determine their precise usage conditions. A more strongly grammaticised form of the possessive article is found in Amharic (cf. Kapeliuk 1994).

4.4. Further (minor) sources

The sources for articles discussed in the preceding sections certainly do not exhaust all possibilities. In at least four instances it is clear that the source of an article in a given language is not a demonstrative, numeral, or personal or possessive pronoun.

Two instances pertain to the honorific and emotional articles mentioned in § 3.2.1 above. No unequivocal sources for these kinds of articles have been identified so far (for one of the Polynesian emotional-diminutive articles (Samoan *si* and its cognates) it has been suggested that it derives from the word for 'little'). Note that these elements are only considered articles if they are in complementary distribution with other kinds of articles (typically NP-articles). Otherwise, they would be considered simply honorific or diminutive particles without further import to the structure of nominal expressions.

The definite article in Sardinian is derived from the Latin identity pronoun *ipse* 'self' (Selig 1992: 177–185) which was frequently used in anaphoric contexts in late Latin (Selig 1992: 117–120) and also survives as a highly general and frequent marker of anaphoric reference in some Catalan dialects (where it is one of the two so-called definite articles, see § 3.2.1 above).

Finally, in Salish there is a determiner based on the root *k^w*-which is generally glossed as 'remote, hypothetical, or conjectured' in the grammars of these languages. It is fairly frequent in nominal expressions and occurs in typical anaphoric contexts:

- (25) Lushootseed (Hilbert & Hess 1977: 19)
tu-lə-?u.ɣ̃w čəd dx^w?al k^wi
 WILL-SERIES-go I TOWARD REMOTE
si?ab ?ab(s)-s-ləɣ̃-il
 noble HAVE-ABS-light-BECOME
 'I will be going towards the noble
 man who possesses the daylight.'

This element also appears to be used as a subordinator (complementizer) with modal implications. It is clear that it is in complementary distribution with the deictic determiners mentioned in § 3.2.1 and hence belongs to the paradigm of determiners. It is also reasonably certain that it is not derived from a numeral, deictic, or pronoun. But its origin and precise meaning and function remain to be determined.

5. Co-occurrence with other determiners

One major typological parameter relevant to all articles irrespective of their source is the question of whether or not articles may co-occur with other determining elements such as possessive pronouns and demonstratives. ‘True’ indefinite articles (as opposed to numerals) generally do not co-occur with any other determiner, for obvious semantic reasons. Specific articles quite generally may co-occur with both possessive pronouns and demonstratives (e.g., Tagalog *ang anak ko* ‘my child (SPEC child 1.SG.POSS)’). In many languages, definite articles also freely co-occur with possessive pronouns as in Italian *il mio amico* ‘my friend (the my friend)’ (cf. Plank 1992). Somewhat less commonly, definite articles may co-occur with demonstratives, for example, in Gulf Arabic (Holes 1990: 175) *haadha l-garaar* ‘this decision (PROX DEF-decision)’ (similar constructions are found in Hungarian, Welsh and Greek). Co-occurrence phenomena such as the precedings ones have been amply documented in recent work by Plank et al. (Plank 1995) under the label ‘overdetermination’. This also includes the co-occurrence of two different forms of an article within the same nominal expression as in the Swedish example *den gamle mann-en* ‘the old man’ mentioned in § 3.2.1.

Acknowledgements

Many thanks for helpful discussion and data to Kumara Henadeerage (Sinhala), Wayan Pastika (Balinese), Andy Pawley (Polynesian), and Heinz Roberg (Nama). All Tagalog and Indonesian data without source references are from the author’s corpus.

6. Special abbreviations

ALL	allative
BEN	benefactive

CMA	complement article
COMPL	completive
DIST	distal demonstrative
INCH	inchoative
INDEF	indefinite
INVIS	invisible
LK	linker (linking article)
MED	medial demonstrative
PM	predicate marker
PROX	proximal demonstrative
REAL	realis
RED	reduplication
REL.DES	relative desinence
RPRT	reportative
SPEC	specific
STAT	stative
VIS	visible

7. References

- Anderson, Stephen R. 1992. *A-morphous morphology*. Cambridge: Cambridge University Press.
- Bechert, Johannes. 1993. *Definiteness and article systems*. Eurotyp Working Papers I/4.
- Bisle-Müller, Hansjörg. 1991. *Artikelwörter im Deutschen*. Tübingen: Niemeyer.
- Bloomfield, Leonard. 1917. *Tagalog texts with grammatical analysis*. Urbana: The University of Illinois.
- Boas, Franz. 1911. “Kwakiutl”. In: Boas, Franz (ed.). *Handbook of American Indian languages*. Washington: Bureau of American Ethnology, I: 425–557.
- Boas, Franz. 1947. *Kwakiutl grammar with a glossary of the suffixes*. Philadelphia: The American Philosophical Society.
- Braunmüller, Kurt. 1982. *Syntaxtypologische Studien zum Germanischen*. Tübingen: Narr.
- Buchholz, Oda & Fiedler, Wilfried. 1987. *Albanische Grammatik*. Leipzig: Verlag Enzyklopädie.
- Chafe, Wallace L. (ed.). 1980. *The Pear Stories: Cognitive, cultural, and linguistic aspects of narrative production*. Norwood, N. J.: Ablex.
- Chesterman, Andrew. 1991. *On definiteness*. Cambridge: Cambridge University Press.
- Christoffersen, Paul. 1939. *The articles. A study of their theory and use in English*. Copenhagen: Munksgaard.
- Churchward, C. Maxwell. 1953. *Tongan grammar*. London: Oxford University Press.
- Clark, Ross. 1974. “On the origin of the Tongan definitive accent”. *The Journal of the Polynesian Society* 83: 103–108.
- Diessel, Holger. 1999. *Demonstratives: form, function and grammaticalization*. Amsterdam: Benjamins.
- Dryer, Matthew. 1989. “Article-Noun Order”. *Chicago Linguistic Society* 25: 83–97.

- Ebert, Karen H. 1970. *Referenz, Sprechsituation und die bestimmten Artikel in einem nordfriesischen Dialekt*. Bräist/Bredstedt: Nordfriisk Instituut.
- Fillmore, Charles J. 1982. "Towards a descriptive framework for spatial deixis". In: Jarvella, Robert J. & Klein, Wolfgang (eds.). *Speech, Place and Action*. Chichester: John Wiley, 31–59.
- Foley, William A. 1980. "Towards a universal typology of the noun phrase". *Studies in Languages* 4: 171–200.
- Gamillscheg, Ernst. 1937. "Zum romanischen Artikel und Possessivpronomen". In: id. *Ausgewählte Aufsätze von Ernst Gamillscheg*. (Supplementheft XV der Zeitschrift für französische Sprache und Literatur) Jena: Wilhelm Gronau.
- Givón, Talmy. 1981. "On the development of the numeral 'one' as an indefinite article". *Folia Linguistica Historica* 2: 35–53.
- Goddard, Cliff. 1985. *A grammar of Yankunytjatjara*. Alice Springs: Institute for Aboriginal Development.
- Grasserie, Raoul de la. 1896. "De l'article". *Mémoires de la Société de Linguistique de Paris* 9.4: 285–322, 381–394.
- Greenberg, Joseph H. 1978. "How does a language acquire gender markers?" In: Greenberg, Joseph H., Ferguson, Charles A. & Moravcsik, Edith (eds.) *Universals of Human Language*. Stanford: Stanford University Press, III: 47–82.
- Haiman, John. 1977. "Connective particles in Hua: an essay on the parts of speech". *Oceanic Linguistics* 16: 53–107.
- Haiman, John. 1978. "Conditionals are topics!". *Language* 54: 564–589.
- Haiman, John. 1980. *Hua: A Papuan language of the Eastern Highlands of New Guinea*. Amsterdam: Benjamins.
- Hartmann, Dietrich. 1982. "Deixis and Anaphora in German Dialects: The semantics and pragmatics of two definite articles in dialectal varieties". In: Weissenborn, Jürgen & Klein, Wolfgang (eds.) 1982. *Here and there*. Amsterdam: Benjamins, 187–207.
- Hawkins, John A. 1978. *Definiteness and indefiniteness*. London: Croom Helm.
- Hawkins, John A. 1991. "On (in)definite articles: implicatures and (un)grammaticality prediction". *Journal of Linguistics* 27: 405–422.
- Heinrichs, Heinrich Matthias. 1954. *Studien zum bestimmten Artikel in den germanischen Sprachen*. Giessen: W. Schmitz.
- Hilbert, Vi & Hess, Thomas. 1977. "Lushootseed". In: Carlson, Barry F. (ed.). *Northwest Coast Texts*. Chicago: The University of Chicago Press, 4–32.
- Himmelmann, Nikolaus P. 1996. "Demonstratives in narrative discourse: a taxonomy of universal uses". In: Fox, Barbara (ed.). *Studies in anaphora*. Amsterdam: Benjamins, 203–252.
- Himmelmann, Nikolaus P. 1997. *Deiktikon, Artikel, Nominalphrase: Zur Emergenz syntaktischer Struktur*. Tübingen: Niemeyer.
- Himmelmann, Nikolaus P. 1998. "Regularity in irregularity: Article use in adpositional phrases". *Linguistic Typology* 2: 315–353.
- Holes, Clive. 1990. *Gulf Arabic*. London: Routledge.
- Jelinek, Eloise & Demers, Richard A. 1994. "Predicates and pronominal arguments in Straits Salish". *Language* 70: 697–736.
- Kapeliuk, Olga. 1994. *Syntax of the noun in Amharic*. Wiesbaden: Harrassowitz.
- Krámský, Jiří. 1972. *The article and the concept of definiteness in language*. The Hague: Mouton.
- Laury, Rita. 1997. *Demonstratives in interaction: The emergence of a definite in Finish*. Amsterdam: Benjamins.
- Lefebvre, Claire & Massam, Diane. 1988. "Haitian Creole syntax: a case for DET as head". *Journal of Pidgin and Creole Languages* 3: 213–243.
- Lewis, G. L. 1967. *Turkish grammar*. Oxford: Clarendon Press.
- Montler, Timothy. 1986. *An outline of the morphology and phonology of Saanich, North Straits Salish*. (University of Montana Occasional Papers in Linguistics No. 4) Missoula: University of Montana.
- Moravcsik, Edith A. 1969. "Determination". *Working Papers on Language Universals* (Stanford) 1: 64–98.
- Mosel, Ulrike & Hovdhaugen, Even. 1992. *Samoan reference grammar*. Oslo: Scandinavian University Press.
- Mulder, Jean G. 1994. *Ergativity in Coast Tsimshian (Sm'algyax)*. Berkeley: University of California Press.
- Plank, Frans. 1992. "Possessives and the distinction between determiners and modifiers (with special reference to German)". *Journal of Linguistics* 28: 453–468.
- Plank, Frans (ed.). 1995. *Overdetermination*. Eurotyp Working Paper (Theme 7, No. 24).
- Reesink, Ger P. 1994. "Domain-creating constructions in Papuan Languages". In: Reesink, Ger P. (ed.) *Topics in Descriptive Papuan Linguistics*. (Semaian 10) Leiden: Vakgroep Talen en Culturen van Zuidoost-Azië en Oceanië, 98–121.
- Scatton, Ernest A. 1984. *A Reference grammar of Modern Bulgarian*. Columbus: Slavica Publishers.
- Selig, Maria. 1992. *Die Entwicklung der Nominaldeterminanten im Spätlatein*. Tübingen: Narr.
- Tung, T'ung-ho. 1964. *A descriptive study of the Tsou Language*. Taipei: Academia Sinica.
- Wilkins, David P. 1989. *Mparntwe Arernte (Aranda): Studies in the structure and semantics of grammar*. PhD thesis ANU, Canberra.

Nikolaus P. Himmelmann,
Ruhr-Universität Bochum (Germany) and
Australian National University, Canberra
(Australia)