Competing Methods for Uncovering Linguistic Diversity: The Case of Definite and Indefinite Articles[1]

Matthew S. Dryer

University at Buffalo

1. Introduction

Davis, Gillon and Matthewson (hereafter Davis et al) argue that "hypothesis-driven fieldwork has led to a more complete and accurate picture of linguistic diversity than has been produced by methods which rely largely on extracting information from existing descriptive grammars". They discuss five cases to support their conclusion. In this response, I will discuss only one of these cases, since this case involves two chapters of mine (Dryer 2005a, 2005b, 2013a, 2013b) in the World Atlas of Language Structures (Haspelmath et al 2005, Dryer and Haspelmath 2013) (hereafter the WALS atlas).[2] These chapters deal with definite and indefinite articles among the languages of the world (though for reasons discussed in §6, I use these labels in a considerably broader sense than many would expect). I will argue that the research for these two chapters uncovered

[1] I am indebted to Juergen Bohnemeyer, Martin Haspelmath, and Lea Brown for very helpful comments on an earlier draft of this paper.

[2] The WALS atlas was originally published as a book (Haspelmath et al 2005) and since 2008 has existed online (http://wals.info/). Davis et al cite the 2011 online edition. The most recent edition is the 2013 edition, which differs in content from the 2011 edition only in correcting some errors. I cite the 2005 book edition as well since, as discussed below, a few properties of my WALS chapters on definite and indefinite articles need to be understood in the context of these chapters originally being part of an atlas in book form.

types of diversity that the methodology that Davis et al employ is unlikely to uncover. For reasons of space, I will only be able to summarize some of this diversity here, which will be described in more detail in Dryer (in preparation).[3]

The primary goal of this response is to illustrate the diversity of articles that has been uncovered by the research behind my WALS chapters on definite and indefinite articles. In §2, I discuss some preliminary issues as well as what appears to be a misunderstanding by Davis et al regarding the methodology behind my WALS chapters. In §3, I discuss a classification of noun phrases in terms of what I call the Reference Hierarchy that is necessary background both for my illustrating the diversity uncovered by the research behind my WALS chapters and for explaining my criteria for defining definite and indefinite articles for the purposes of my WALS chapters. In §4, I present a basic typology of articles, based on the Reference Hierarchy. This provides a way of summarizing some of the diversity uncovered by my research. In §5, I briefly discuss some more unusual articles, providing more evidence of the diversity my methodology uncovered. In §6, I explain in detail my use of the terms *definite* and *indefinite* for the purposes of my WALS chapters, and why, by my criteria, my classification of articles in the Salish languages appears to be correct. In §7, I discuss the general issue of uncovering diversity, arguing that while there are clear limitations on what my methodology can uncover, there are some aspects of diversity that it is well-suited for. In §8, I discuss two general methodological issues that Davis et al discuss at length, namely

[3] My interest in definiteness dates back to Dryer (1973), which was primarily about definiteness in English.

the use of falsifiable hypotheses and the need for negative evidence, and argue that in neither case is there much disagreement. First, I argue that my work also employs falsifiable hypotheses. Second, I concede that my methodology is limited in what it can accomplish precisely because of the lack of negative evidence, but argue that the evidence presented in the earlier sections illustrates how my methodology is well-suited for uncovering certain kinds of diversity. My conclusion is that the optimal way to uncovering diversity is the use of different methodologies, that there is no unique way to pursue diversity.

I need to be clear from the start that I in no way question the value of the work that Davis, Gillon, and Matthewson have done in this area. There is no question that there are certain kinds of diversity that detailed examination of a single language can uncover that the methodology I used in these studies, based on extracting information from existing descriptive grammars, cannot uncover. In fact, for few of the over 300 languages I examined for these two WALS chapters of mine are the factors conditioning the use of articles as clear as those governing the use of determiners in Skwxwú7mesh by Gillon (2006) and St'át'imcets by Matthewson (1998), other than a number of European languages where the use of definite and indefinite articles is fairly well-known independently of the research I did. Furthermore, their work clearly has value in documenting languages, quite apart from issues of uncovering diversity. In other words, the issue is not the relative value of their research and mine, but simply the more specific question of how good a particular methodology is in uncovering diversity.

2. Davis et al's claim's about my methodology

Davis et al's argument that my methodology failed to uncover diversity in the use of articles is based on two misunderstandings. The first is one that I have to accept full responsibility for. The authors of WALS chapters had only two pages in which to discuss various aspects of their typology and I was forced to discuss my criteria for identifying definite and indefinite articles rather briefly. What I did say covered most languages, but there was not sufficient space to lay out the full criteria (described in §6) that would explain my classification of all the languages I have examined. Unfortunately, one of the few languages that would have required such a detailed explanation of my criteria was the Salish language Sḵwx̱wú7mesh, the language that Davis et al say I misclassified. Once I have outlined my methodology below, I will explain how my classification of Sḵwx̱wú7mesh is in fact correct. However, nobody can fault Davis et al for concluding that I have misclassified them. Furthermore, I do not want to suggest that every instance of how I classify languages is accurate. There is little question that I have probably misclassified many languages, either because my source is inaccurate or misleading or because I have misinterpreted the source. I have to concede to Davis et al that the methodology I use, of collecting data from grammars for a large number of languages, runs the risk of many instances of misclassification.[4]

[4] Davis et al suggest that the accuracy of WALS would be increased if it allowed experts on particular languages to be able to change coding of languages in a wiki manner, as is the case with SSWL (The Syntactic Structures of the World Languages). While I believe that SSWL provides a valuable complement to WALS and will hopefully become more important through time, our experience with WALS is that the vast majority of alleged errors brought to our attention by experts are not in fact errors, that the experts have not

The other misunderstanding by Davis et al is that they do not seem to understand the process in collecting data of the sort I collected for these chapters. What I have done involves two activities. The first is the collection of data on articles from a large sample of languages, which has uncovered the diversity I describe below. The second activity involves dividing the multi-dimensional typological space representing differences among languages into regions and assigning languages to a number of types. Furthermore, because WALS is an atlas and the editors felt that the goals of the atlas would be better served by maps showing a smaller number of feature values, authors were asked to classify languages into a smaller set of types than might otherwise be used. In other words, the particular way I divided the typological space was motivated in part by what served the goals of an atlas. But what I refer to as my methodology in this paper is what was involved in the first activity, where I have uncovered the diversity to be summarized below.

Davis et al's conclusion that I have misclassified a number of Salish languages largely stems from the fact that I use the terms *definite* and *indefinite* in a way that is considerably looser than the way these terms are generally used. But this broad usage reflects what is the most important result of my study. Namely, languages with an obligatory binary contrast between definite and indefinite articles of the sort found in

---

read the defining criteria carefully. Thus allowing experts to make changes would result, in the majority of cases, to changing correct codings to incorrect ones and would thus decrease the accuracy, not increase it.

English are actually relatively uncommon outside of Europe (and the Middle East).[5] Most languages that have articles that code meanings within the general semantic domain of definiteness or indefiniteness (in a sense to be explained in §4) exhibit a distribution that is clearly different from that of the definite and indefinite articles in English and many other European languages. In particular, most languages that have articles that are restricted to what would normally be considered definite contexts use these articles in only a subset of the contexts in which English would use a definite article. My data on this is not organized in a way that makes it easy to cite exact numbers, but in so far as can estimate, I would say that in perhaps only about 10% to 20% of languages with articles that are restricted to definite contexts do these articles occur in most of the contexts where English would require a definite article. Examples of languages where there is an article that is restricted to definite contexts but is not obligatory in such contexts, excluding instances of articles where my source gives some indication how they are restricted, include Busa (Wedekind 1972), Mongsen Ao (Coupe 2007), Loniu (Hamel 1985), Alamblak (Bruce 1984), Maricopa (Gordon 1986), Chontal Maya (Knowles 1984), and Lokono Dian Arawak (Pet 1987). [6]

Similar comments apply to indefinites. In other words, most languages that have two articles, one which is restricted to definites and one that is restricted to indefinites,

[5] Note that for a language to have an obligatory definite article does not exclude the possibility that the definite article is absent from noun phrases which contain other indications of the definiteness of the noun phrase, like a demonstrative or a possessor.
[6] Note that some of the languages and sources I cite in this paper are not reflected in my current WALS chapters; all my current WALS chapters simply reflect a stage in an ongoing project.

distinguish three categories formally: those marked definite, those marked indefinite, and those not marked as definite or indefinite. And in most languages that have one article, say an article that is restricted to definites, this article does not occur on all definite noun phrases, so the absence of an article does not mean that the noun phrase is indefinite.

For the purposes of this response and of my WALS chapters on definite and indefinite articles, I use the term 'article' in a somewhat nonstandard sense, one that is more semantic than common uses of the term. Specifically, although I restrict it to words or morphemes that occur in noun phrases, I otherwise make no assumptions as to the syntactic status of the words or morphemes in question. Namely they must code something in the general semantic domain of definiteness or indefiniteness (in a sense to be explained in §4). In some languages, they appear to be affixes rather than separate words and thus are not articles (or determiners) in any syntactic sense.

I do, however, exclude other sorts of words that are inherently definite because of the nature of their more specific meaning. The most obvious type of word of this sort is demonstratives. My criteria for whether something is a demonstrative is that it have one of two properties: (1) it involves some sort of contrast in spatial distance, usually contrasting proximal with one or more spatial distal categories; or (2) its basic use is exophoric, referring to something in the shared perceptual space of the speaker and hearer. However, if one of the demonstratives (almost always a distal one) is widely used anaphorically, I do treat that as a definite article, though in my WALS chapter I separate these out from definite articles that are distinct in form from the demonstratives. I thus

do not count something as a definite article if it involves some sort of distal contrast, unless the distal constrast is used or can be used to distinguish two anaphoric uses, where one is used for very recent mentions, the other for mentions further back in the discourse (see more on this below). This is particularly relevant to my response to Davis et al, since I do not classify some Salish languages as having definite articles because the words in question appear to involve some sort of spatial distal contrast.

My notion of article appears to be similar to Davis et al's notion of determiner. I do not use the term 'determiner' since for many linguists that term is understood to include demonstratives as well. Davis et al's notion of determiner ("we assume that determiners are elements that introduce argument noun phrases, and cannot occur on their own (in contrast to demonstratives)") differs from my notion of article in three respects. First, my notion of article excludes adnominal demonstratives which are different in form from pronominal demonstratives (Diessel 2013) and which therefore cannot occur on their own. Second, my notion includes adnominal words coding something like definiteness that also function as third person pronouns (Dryer 1989). And third, my notion of article includes words that do not introduce argument noun phrases, but occur elsewhere in the noun phrase, most commonly the end of the noun phrase, as in Abui (Kratochvil 2007), Busa (Wedekind 1972) or Maricopa (Gordon 1986) (though I suspect that Davis et al did not really intend to exclude words that do not occur at the beginning of noun phrases).

3. The Reference Hierarchy

The primary goals of this paper are to illustrate the diversity that has bee uncovered by my research on articles and to explain the criteria I use for identifying definite and indefinite articles for my WALS chapters. However, before I can do this, I must first discuss a classification of noun phrases that will provide the basis for a typology of articles. This classification involves five types of noun phrases that are ordered in what I call the Reference Hierarchy.

The Reference Hierarchy

anaphoric definites > nonanaphoric definites > pragmatically specific indefinites
> pragmatically nonspecific (but semantically specific) indefinites
> semantically nonspecific indefinites

The primary factor defining the ordering of these types on the hierarchy is the following: if a language has an article that is used for more than one type of noun phrase on the hierarchy, then the set of types it is used with will be a set that is contiguous on the hierarchy. For example, the definite article is used in English with noun phrases of the first two types on the hierarchy while the indefinite article in English is used with noun phrases of the last three types. A second factor is an underlying semantic or pragmatic notion that one might characterize informally as running from “most definite” to “least definite” that explains why articles are only used for contiguous sets of types on the

hierarchy.

The Reference Hierarchy is an adaptation of a similar hierarchy (or more accurately 'wheel') proposed by Givón (1978).[7] Ultimately, a more complete classification of noun phrases would have to include generics and true predicate nominals, that is nominals which function semantically as predicates, in contrast to nominal predicates in equational clauses, the former represented in English by indefinite noun phrases in predicate position (e.g. *He is a teacher*), the latter by definite noun phrases in predicate position (e.g. *He is the teacher*). True predicate nominals are most commonly coded the same way as semantically nonspecific indefinites while generics are coded like either semantically nonspecific indefinites or nonanaphoric definites (cf. Givón 1978, 1984: 407). The Reference Hierarchy also does not include pronouns, since it is intended as the basis of a typology of articles, though it is clear that if pronouns were included, they would constitute the highest position on the hierarchy.

A detailed description of the five types of noun phrases in the Reference Hierarchy is not possible here, but they have all been discussed elsewhere in the literature, though not always with the same labels. The first position on the hierarchy, anaphoric definites, covers definite noun phrases that refer back in the discourse while the second

[7] My Reference Hierarchy differs from Givón's (1978) wheel of reference in distinguishing anaphoric definites from nonanaphoric definites and in not including generics and predicate nominals. I also use different terminology from Givón: my 'pragmatically specific indefinite' corresponds to his 'referential indefinite', my 'pragmatically nonspecific but semantically specific indefinite' corresponds to his 'referential nondefinite', and my 'semantically nonspecific' correspondings to his 'nonreferential object', except that my notion is not restricted to (grammatical) objects.

position, nonanaphoric definites, covers definite noun phrases that do not, where their use is based only on shared knowledge of the speaker and hearer; the notion of anaphoric definite is close to the notion of discourse-old in the sense of Prince (1992). However, previous mention in a text does not necessarily make a definite noun phrase an instance of an anaphoric definite. For example, in languages with an article specifically marking anaphoric definites, this marker would not occur with noun phrases denoting the sun, even when there is a previous reference in the text to the sun, just as in English, one would not normally refer to the sun with the noun phrase *the aforementioned sun*, even if there were a previous reference to the sun; we would only do so in unusual contexts where multiple suns were involved. In other words, anaphoric definites are ones whose use is licensed by linguistic antecedents. My distinguishing anaphoric from nonanaphoric definites is motivated largely by the fact that articles coding anaphoric definites are very common among the world's languages, certainly more common than articles coding pragmatically nonspecific indefinites or articles coding semantically nonspecific indefinites and probably more common than articles with a distribution close to that of either the definite article or the indefinite article in English. Examples of languages with articles restricted to anaphoric definites include Garrwa (Mushin 2012), Tariana (Aikhenvald 2003), Lamang (Wolff 1983), Lampung (Walker 1976), and Araki (François 2002).

The three types of indefinites, the last three types on the Reference Hierarchy, are distinguished by two quite distinct notions of specificity, namely pragmatic specificity and semantic specificity. The distinction between semantically specific and semantically

nonspecific is what is probably the more familiar notion of specificity, where semantically specific noun phrases are associated with an entailment of existence (e.g., *John bought a car* entails that a car exists) and semantically nonspecific noun phrases are not (as in *John is looking for a unicorn*). The distinction between pragmatic specificity and pragmatic nonspecificity is less well-known, but often recognized (e.g., Givón 1978, Fodor and Sag 1982, Gernsbacher and Shroyer 1989, Enç 1991, Payne 1992: 150, Gundel et al 1993: 276, Schroeder 1999: 56). It is difficult to define precisely, but strongly correlates with subsequent reference: a pragmatically specific indefinite noun phrase normally introduces a participant into the discourse that is referred to again in the subsequent discourse, while a pragmatically nonspecific indefinite noun phrase normally does not. For example, if someone says *I went to a movie last night*, and then goes on to talk about the movie, then *a movie* is normally pragmatically specific. However, if someone says the same sentence but then proceeds to say nothing more about the particular movie, then *a movie* is normally pragmatically nonspecific. Note that in both cases here, *a movie* is semantically specific, since in both cases there is an entailment that there is a movie. The use of the word *this* with indefinites in English is an example of an indefinite marker that is restricted to pragmatically specific indefinites (Prince 1981, Gernsbacher and Shroyer 1989, Gundel et al 1993): if one starts a conversation with *I went to this movie last night*, there is an expectation that one will make subsequent reference to the particular movie.

The notion of pragmatic specificity is probably the least well-known of the five notions on the Reference Hierarchy. But the motivation for including it is that the use of

articles in many languages is clearly sensitive to such a notion; in fact such articles are probably more common among the languages of the world than articles with a distribution like the indefinite article in English which occur with most indefinite noun phrases (or most singular indefinite noun phrases with count nouns). In other words, there are many languages with articles that are restricted to a subset of semantically specific indefinites, where what characterizes this subset is some sort of prominence in the discourse, a prominence that correlates strongly with subsequent mention.

Note that the Reference Hierarchy assumes that semantically nonspecific noun phrases are necessarily pragmatically nonspecific (even though it is possible to continue referring to something that is introduced with a semantically nonspecific noun phrase: *John is looking for a new house. It must be in the city and it must be at least 80 years old.*). The basis for this assumption is that articles that code pragmatic specificity appear never to occur with semantically nonspecific noun phrases.[8]

4. A basic typology of articles

In this section, I lay out a basic typology of articles that emerges from examining

[8] There are some instances of formally definite noun phrases in English that might be considered nonspecific. For example, some instances of weak definites (Carlson and Sussman 2005), such as *the radio* in *Everyone heard about it on the radio,* share properties with nonspecific noun phrases. However, I believe, on the basis of what I have observed about how definite articles are used in texts for other languages, that such noun phrases would rarely be marked as definite in other languages. I suspect that this use of the definite article in English is related to the fact that English can use the definite article to mark generic noun phrases and I would predict that we would never find noun phrases like this marked definite in languages that do not use the definite article with generics.

articles in many languages, a typology that is defined in terms of the Reference Hierarchy. This has two purposes. One is that it illustrates the diversity that the research for my WALS chapters uncovered. The other that this typology is necessary background for explaining how I use the terms *definite* and *indefinite* for the purposes of my WALS chapters, as discussed in §6 below.

As noted above, the primary factor defining the order on the Reference Hierarchy is the claim that if an article in some language can be used with more than one of the five types on the Reference Hierarchy, then the set of types it can be used with will be contiguous on the Reference Hierarchy. What this means is that the set of possible articles is claimed to be those given in Table 1.

Table 1 should be interpreted as follows. Each of the columns represents one of the five types of noun phrases represented in the Reference Hierarchy.[9] Each of the rows represents one of the types of articles that is possible if the set of noun phrase types that an article can occur with must be a contiguous set of types in the Reference Hierarchy.[10] For each type of article, the thick black lines represent the section of the Reference

---

[9] I abbreviate the types of noun phrases as 'AD' for anaphoric definites, 'ND' for nonanaphoric definites, 'PSI' for pragmatically specific indefinites, 'PNI' for pragmatically nonspecific (but semantically specific) indefinites, and 'SNI' for semantically nonspecific indefinites. Using these abbreviations, the Reference Hierarchy can be restated as AD > ND > PSI > PNI > SNI.

[10] Informal characterizations of some of the types of articles, corresponding to the lines in Table 1, are definite for AD+ND, semantically specific indefinite for PSI+PNI, pragmatically nonspecific for PNI+SNI, pragmatically specific for AD+ND+PSI, indefinite for PSI+PNI+SNI, semantically specific for AD+ND+PSI+PNI, and discourse-new (or nonanaphoric) for ND+PSI+PNI+SNI.

Hierarchy which that article occurs with. An example of a language with an article of that type is given above the thick black line.[11]

[11] My sources for the languages mentioned in Table 1 are Mushin (2012) for Garrwa, Blackings and Fabb (2003) for Ma'di, Smye (2004) for Anufo, Samarin (1966) for Gbeya Bossangoa, Schuh (1972) for Ngizim, Ross (2002) for Siar, Palmer (2009) for Kokota, Hooper (1993,1996) for Tokelauan, Dayley (1985) for Tzutujil, and Hualde and de Urbina (2003) for Basque.

type of article \ type of noun phrase	anaphoric definite (AD)	nonanaphoric definite (ND)	pragmatically specific indefinite (PSI)	pragmatically nonspecific but semantically specific indefinite (PNI)	semantically nonspecific indefinite (SNI)
AD	Garrwa				
ND		Ma'di			
PSI			Anufo		
PNI				unattested	Gbeya Bossangoa
SNI					
AD+ND	English (def)				
ND+PSI		unattested			
PSI+PNI			Ngizim		
PNI+SNI				Siar	
AD+ND+PSI	Kokota				
ND+PSI+PNI			unattested		
PSI+PNI+SNI				English (indef)	
AD+ND+PSI+PNI	Tokelauan				
ND+PSI+PNI+SNI			Tzutujil		
all five types			Basque		

Table 1
A Preliminary Typology of Articles

A number of comments are in order. The first and most important one is that each of the types in Table 1 includes both articles which are obligatory for the type in question and articles which are restricted to the type in question but are not obligatory for that type. There are two ways in which this situation might arise, two ways in which an article

might be restricted to a particular type but are not obligatory for that type. First there are situations where there seems to be some definable subset of cases within a type that are treated differently from other cases; in these situations the meaning of the article will be narrower than that of the larger type. An example of this, to be mentioned below in §5, is an article for anaphoric definites where the previous mention is not in the immediately preceding discourse but further back. Second, there may be situations where the use of an article within a type seems to be truly optional in the sense that the apparent meaning of the article is as broad as the type, but the use versus nonuse is conditioned by intended perlocutionary effect (Austin 1962).[12] The difference between speaking with a normal volume and speaking more loudly in a particular situation because one is concerned that the hearer may not understand the intended meaning is not a difference in meaning, but a difference in intended perlocutionary effect. Choices conditioned by intended perlocutionary effect appear to be common with many phenomena in language, as with optional case markers, where the case marker is more likely to be used if there is a greater risk of ambiguity or misunderstanding if it is not used. The factors conditioning such optional uses of morphemes are not any more a matter of meaning than speaking more slowly or speaking more loudly. A language with a truly optional definite article would be one where the article can be used for any type of definite noun phrase, but is only used if the speaker is concerned that the intended meaning might not be conveyed if the article were not used.

[12] There are also probably cases where both situations described here obtain, articles whose meaning is narrower than that of the broader type, but whose use is still truly optional even within the narrower meaning.

Unfortunately, in most cases where it is clear that an article is not obligatory within a type, it is not clear which of the two possibilities just described obtains. In a few cases, the narrower use is described by an author, as in the special uses of articles described in §5 below. However, in most cases, it is not clear what governs the use of articles that are not obligatory within their type, whether they have narrower meaning or are truly optional. Furthermore, the vast majority of articles I have examined are in fact not obligatory within their type. Sometimes authors will say that an article is not obligatory but without further explanation and sometimes it is only clear from examples that it is not.

It is also the case that there are many languages where my source characterizes something as an indefinite article but where all that one can determine from the source is that it is restricted to indefinites but not obligatory for indefinites. In such cases, it is possible that it is restricted to pragmatically specific indefinites or to semantically specific indefinites, but one cannot tell. It is not possible to classify these languages into one of the types in Table 1. And one finds an analogous situation with articles in some languages that are described as definite articles, but where these articles are not obligatory with definite noun phrases. In fact, the majority of articles that I have examined cannot be classified into one of the types in Table 1, because of insufficient data.

One might question the inclusion of the last type in Table 1, articles which can be used with all five types of noun phrases on the hierarchy. And certainly if an article is

obligatory for all types on the Reference Hierarchy, then I would exclude it from consideration since my focus is on articles that have meanings related to notions like those on the Reference Hierarchy. My database includes data on articles that do not code anything related to the Reference Hierarchy (such as articles in Austronesian languages which code only the distinction between common nouns and proper nouns), but these were not considered for my WALS chapters on definite and indefinite articles. However, there are languages like Basque, where there is an article usually called a definite article that can occur with any of the five types except for a subset of semantically nonspecific noun phrases (Hualde and de Urbina 2003: 119-120).[13]

There are three types in Table 1 that are not attested in my data. It is not clear whether these are accidental gaps. The first unattested type is articles that specifically code pragmatically nonspecific but semantically specific indefinites. The absence of such articles in my data reflects a more general property of articles in my data: I am not aware of any language where both notions of specificity are relevant to articles in the language, where the distinction between pragmatically specific and pragmatically nonspecific is relevant to the use of one article, but where the distinction between semantically specific and semantically nonspecific is relevant to the use of another article.

[13] When I characterize articles with meanings in the general semantic domain of definiteness and indefiniteness, what I mean is that their meaning conveys something in terms of the notions in the Reference Hierarchy. For example, the so-called definite article in Basque cannot be used for a subset of semantically nonspecific indefinites, namely those within the scope of negation, and thus has a meaning in the general semantic domain of definiteness and indefiniteness (Hualde and Urbina 2003). In other words, the Basque definite article, when used within the syntactic scope of negation, must be interpreted as semantically specific.

The other two unattested types both involve combinations of nonanaphoric definites with different subsets of indefinites, but ones which do not include anaphoric definites. In fact, apart from articles that are used only with nonanaphoric definite noun phrases, my data includes only one language with an article that is used with nonanaphoric definites but not with anaphoric definites, namely Tzutujil.[14] Furthermore in all the languages in my data where there is an article used specifically for nonanaphoric definites, there is another article that is either for specifically anaphoric definites (e.g. Lampung; Walker 1976) or for definites in general (e.g. Araki; François 2002), where in the latter case nonanaphoric definites occur with either the nonanaphoric definite article or the general definite article and it is not clear what conditions this choice.

The set of types in Table 1 that are attested represents one example of the sort of diversity that my research for my WALS chapters uncovered; it not clear from the work of Davis et al whether they are aware of this particular sort of diversity.

[14] The characterization of Tzutujil by Dayley (1985: 255) implies that what he calls the indefinite article can be used for nonanaphoric definites as well as at least some indefinites, but it is not clear whether it can be used for all three types of indefinites, so it is not clear whether it is specifically ND+PSI+PNI+SNI rather than ND+PSI or ND+PSI+PNI. Furthermore Tzutujil does appear to distinguish nonanaphoric definites from indefinites in that there is a definite article that occurs alone with anaphoric definites but combines with what Dayley calls the indefinite article to mark what appears to be nonanaphoric definites.

5. Special uses of articles

The typology in Table 1 summarizes the basic findings of the research behind my WALS chapters on definite and indefinite articles, although I believe that the most important finding of this research is the fact that articles in most languages are not obligatory within any of these types. However, there is additional diversity that I have uncovered that involves idiosyncratic articles in various languages. In this section, I describe six of these cases briefly. All six involve anaphoric definites.

a. In addition to languages with articles that specifically code anaphoric definites, there are languages that have more than one such article, where a distinction is made between more recent mentions and more distant mentions in the same discourse (e.g. Ma'di (Blackings and Fabb 2003)).

b. In Yapese (Ballantyne 2005), there is an article that codes anaphoric definites but which can also be used for entities that have not been explicitly mentioned but that are highly accessible in the discourse context, in the sense of being on the edge of the hearer's consciousness, entitiess that are inferrable (in the sense of Prince 1992) from other entities that have been mentioned in the preceding discourse, such as *the door* in Prince's example *He passed by the Bastille and the door was painted purple*. This article in Yapese apparently cannot be used with nonanaphoric definites which are not highly accessible in this sense.

c. In Oksapmin (Loughnane 2009), there is an article that marks only those anaphoric definites whose referent was introduced into the discourse only fairly recently. (This article contrasts with an article for nonanaphoric definites, an article that appears to be possible with any definite noun phrase, as well as an indefinite article.)

d. In Biak (van den Heuvel 2006), there is an article that can be used for anaphoric definites and some nonanaphoric definites, but not for referents which are inherently unique and which are known to everyone, like the moon.

e. In Epena Pedee (Harms 1994), there are two articles for anaphoric definites, one marking referents that are particularly important in the discourse and one that appears to be possible with anaphoric definites in general.

f. In Abui (Kratochvil 2007), there are two articles for anaphoric definites, one for referents that the speaker mentioned and one for referents that the hearer mentioned.

6. Defining definite and indefinite articles for my WALS chapters

It is worth considering what a typology which was more fine-grained than the typology in Table 1 would look like. First, for each of the types in Table 1, it would further distinguish those articles which are obligatory within the type and those which are not. Second, it would further distinguish those which are not obligatory because they code some narrower meaning from those that are truly optional (in the sense explained above). And third, among articles that code some narrower meaning, it would distinguish different articles on the basis of what that narrower meaning is. However, only a small minority of the languages I have examined can be classified by such a fine-grained typology, because of lack of data. In fact only a minority of the languages I have examined can be classified by the broader typology in Table 1. For many languages, one can only get a rather vague idea of what governs the use of articles.

Faced with the vague characterization of the use of articles in many sources, I developed a very broad typology, with very broad notions of definiteness and indefinite, one that has allowed me to classify articles in most languages that have them as definite or indefinite. Namely, if an article is more strongly associated with positions higher on the Reference Hierarchy, I treat it as a definite article while if it is more strongly associated with positions lower on the hierarchy, I treat it as an indefinite article. More specifically, I classify articles as follows: (1) if an article is restricted to the first two types on the Reference Hierarchy (including cases where it is restricted to one of these two types), then it counts as a definite article; (2) if an article is restricted to the last three

types on the Reference Hierarchy, then it counts as an indefinite article; (3) if a language has two articles where in some contexts, one article is more likely to be interpreted as definite, the other article as indefinite, then the first article counts as a definite article and the second counts as an indefinite article; (4) if an article is restricted to a set of types that includes anaphoric definites but does not include semantically nonspecific indefinites or some subset of semantically nonspecific indefinites, then it counts as a definite article. Examples of this fourth possibility include (1) pragmatically specific articles, i.e. articles that occur only with definites and pragmatically specific indefinites; (2) semantically specific articles, i.e. articles that occur with anything but semantically nonspecific indefinites; and (3) articles that can occur with anything except for a subset of semantically nonspecific indefinites. An apparent example of this third case is what is commonly called the definite article in Basque. Hualde and Urbina (2003) report that the definite article in Basque can be used for indefinites, even including indefinites within the scope of an intensional operator like 'want', but not for indefinites within the scope of negation.

A second apparent example of this last possibility, of a determiner that can be used for all five types of noun phrases on the Reference Hierarchy, but with only a subset of semantically nonspecific indefinites, is the determiner *ta* in S<u>k</u>w<u>x</u>wú7mesh, one of the Salish languages that is the focus of Davis et al's discussion. As they point out, not only can this determiner be used with definites and semantically specific indefinites, but also with some semantically nonspecific indefinites, as in their (63) glossed 'Did you buy any fish?'. My reasons for classifying *ta* in S<u>k</u>w<u>x</u>wú7mesh as a definite article and for

classifying a second article *kºi* as an indefinite article are the following.[15] Kuipers (1967: 138) cites contrasts such as "*sa ˈt-šit-ka ta sta ˈqº* ... 'give him the water' with the article *ta* and *sa ˈt-šit-ka kºi sta ˈqº* 'give him (some) water' with the article *kºi* to illustrate the difference between the two articles. This contrast alone implies that there are contexts in which a noun phrase with *ta* is more likely to be interpreted as definite and *kºi* is more likely to be interpreted as indefinite. Further evidence for this is provided by discussion in Gillon (2006). She notes that although *kºi* is not restricted to semantically nonspecific indefinites, only a nonspecific reading is possible in the context of a variety of scopal operators, such as negation and the universal quantifier (p. 119). The fact that only a nonspecific reading is possible in these contexts counts as evidence that *kºi* is more strongly associated with positions lower on the hierarchy. Further evidence that *ta* is more strongly associated with positions higher on the hierarchy is provided by two additional facts reported by Gillon (2006). First, although it can be used with narrow scope readings with respect to negation, it can only be used with wide scope readings with respect to other scopal operators, like universal quantification (p. 95). Second, Gillon notes (p. 97) that in the context given in (1) *ta míxalh* 'the bear' can only be interpreted as referring back to the bear mentioned in the preceding sentence.

(1) ... S-en men kw'élash-t ta míxalh.
nom-1sg.sbj just shoot-tr det bear
'[I went hunting. I saw a bear.] I shot the bear.'

[15] Gillon (2006) represents the article which Kuipers represents as *kºi* as *kwi*.

Even though *ta* can be used in certain nonspecific contexts, the fact that it cannot be used in certain other nonspecific contexts and the fact that it must be interpreted as definite in contexts like that in (1) provide sufficient evidence that it counts as a definite article in my sense.

Whether I have correctly classified Bella Coola and Comox, the two other Salish languages discussed, is less clear, but part of the issue for me is whether these languages have determiners that are neutral with respect to distance; I treat words that code a spatial distal contrast as demonstratives, not articles. For example, I did not count the Salish language Upriver Halkomelem as having a definite article since there appears to be a distal contrast with the relevant forms. My interpretation of the discussion of Bella Coola by Davis and Saunders' (1997: 87-88) is that there are forms that are neutral with respect to distance. However, this is not obvious from their discussion, and if what Davis et al say about Bella Coola is correct, then I have indeed misclassified it, even by my own criteria. With respect to Comox the following quotation from Davis et al provides sufficient evidence that it has a definite article in my sense: "Hagège actually states that the Comox system is *not* constructed according to a definite/indefinite opposition, although he writes that the articles he glosses as 'anaphoric, distal' are often used as indefinites, while the articles glossed as 'anaphoric, proximal' are often used as definites". The two other Salish languages that Davis et al discuss, St'át'imcets and Nsyílxcen, are not ones I have coded for this feature in my database.

As noted above, I was not able to fully explain my notion of definiteness and

indefiniteness in my WALS chapters for reasons of space. What I said covered most languages. Unfortunately, Sk̲wx̲wú7mesh was one of the few languages not covered by my brief explanation, where the lengthy explanation that has required much of this response would have been necessary. Davis et al's conclusion that I have misclassified Sk̲wx̲wú7mesh was thus perfectly justified on the basis of what I did say. Although I have to plead guilty to not being clear about my criteria, I believe I have classified Sk̲wx̲wú7mesh correctly according to my criteria.

One might still object, however, to my criteria for determining whether something is a definite or indefinite article. There are actually (at least) two kinds of possible objections. The first objection might be that if I am going to classify languages according to these criteria, I ought to use terms other than 'definite' and 'indefinite'. There is a general issue in typology as to whether to use familiar terms with a meaning somewhat different from their familiar use or new terms. The problem with using familiar terms is that they can have misleading implications; this particular instance could be taken as a good example of that. The problem with new terms is that people often do not understand the general idea behind claims using new terms. In this case, since most of the articles I classified as definite were specifically restricted to definites and most of the articles I classified as indefinite were specifically restricted to indefinites, I felt using the familiar terms was the better option, especially because I did not have space to explain my criteria. But I am aware that there is room for disagreement as to whether my choice of terms was a good one. I should note, however, that my use is not inconsistent with the use by authors of grammatical descriptions. As noted above, what is generally called the

definite article in Basque can be used for indefinites, even some semantically nonspecific indefinites. Similarly, Kuipers (1967) calls the article *ta* in Sḵwx̱wú7mesh a definite article, although it exhibits a somewhat similar distribution to the Basque definite article, at least in terms of the types on the Reference Hierarchy. When authors call something a definite article, they often do not mean that it has a distribution close to that of the definite article in English, just that it has a distribution that is more like that of the definite article than the indefinite article. This is similar to the way I use the terms in classifying languages for my WALS chapters.

The other possible objection is that, apart from my choice of terms, there is a problem with the concepts that I apply the expressions definite article and indefinite article to. Davis et al might object that lumping together everything I treat as a definite article is a mistake. I should emphasize that I do not consider this notion of definiteness particularly useful beyond the needs of my WALS chapter; my notions of definite article and indefinite article are motivated largely because it is not easy to classify languages using more refined notions, because of lack of data.[16] But it is also worth asking what the alternative might be. The discussion by Davis et al (as well as more detailed discussion by Matthewson (1998)) might suggest an alternative typology that includes six types of

[16] Note that apart from my nonstandard use of these terms in my WALS chapters, I otherwise use these terms in a fairly standard sense in this paper, as in the characterization of the typology in Table 1. Note also that my notions of definite article and indefinite article are not intended as crosslinguistic categories; rather, they are comparative concepts in the sense of Haspelmath (2010).

articles: definite, indefinite, specific, nonspecific, indefinite specific[17], and other (where the Salish articles would clearly fall into this last class). One reason why I did not adopt such a typology is the practical one mentioned above: I wanted a typology that would allow me to classify as many languages as possible and for the majority of languages I have examined, there is not clear evidence how to classify them by this typology. One cannot assume from the fact that an article in some language is called a definite article that it is that (rather than, for example, being a specific article). For example Saltarelli et al (1988) refer to the so-called definite article in Basque as such, without apparently noting that it can be used with many types of indefinites.

Furthermore, if the objection to my typology is that it mixes together articles which are very different from each other, then one can make exactly the same objection to the above alternative typology. As discussed above, there are (at least) two different notions of specificity, namely pragmatic specificity and semantic specificity. Despite the choice of labels, these differ from each other as much as they differ from definiteness. It is not clear what justification there might be for grouping these two notions together but distinguishing them from definiteness. Similarly, the difference between anaphoric definite articles and articles that can be used for both anaphoric definites and

[17] One would need to distinguish specific from indefinite specific since there are articles that code specific in general (including definite) and articles that only mark specific indefinites. There is otherwise no need to consider other combinations of definiteness and specificity since definites are always specific (and nonspecifics are always indefinite). The discussion by Davis et al and by Matthewson (1998) seems to assume that determiners in a given language will code either definiteness or specificity, which would leave out articles that code specific indefinites; however, perhaps this was not their intention.

nonanaphoric definiteness is in some ways analogous to the distinction between definite articles and specific articles, in the sense that anaphoric definite noun phrases are a subset of definite noun phrases and definite noun phrases are a subset of specific noun phrases. And given the lack of information on what governs the choice of definite articles in languages where they do not occur on all definite noun phrases, there may be a great variety of different subsets of definites that would be grouped together on the alternative typology.

7. Uncovering diversity

I have attempted in this paper to demonstrate a number of instances of diversity that the research for my WALS chapters on definite and indefinite articles has uncovered. However, I would not want to argue that the methodology behind my WALS chapters is the only way or even the best way to uncover diversity. For one thing, the fact that for most languages, the precise meaning of the article is unclear means that the picture that emerges is "fuzzy" in the sense that it is like a satellite photo of an area of the surface of the earth, where zooming in on an area eventually leads to a blurry picture (as is the case for Google Earth for many parts of the world). A clearer picture would require detailed examination of articles in many different languages.

What the methodology behind my WALS chapters does allow is to give a bird's eye view of the landscape. It is not clear from the work of Davis et al that they are aware of the sort of diversity my study uncovered. For example, it is not clear that they are

aware that languages with articles with a distribution similar to that of the definite article and indefinite article in English are relatively uncommon outside Europe, that in most languages with an article that is restricted to definite noun phrases, the article is used in fewer contexts than in English, and it is not clear that they are aware of the sort of special cases discussed above in §5. In other words, to get a clear picture of the diversity found among the world's languages, we need both crosslinguistic studies of the sort I have conducted and detailed examination of particular languages.[18]

It is also not clear how much further one can go applying my methodology to articles, given the extent to which grammatical descriptions are so often vague as to how articles in particular languages are used. At most, examining more languages might uncover more types and find additional instances of types I have already found. But getting a clearer picture of exactly how languages differ in their articles requires detailed examination of particular languages.

On the other hand, there are questions that can only be answered by the methodology of collecting data from a large number of grammatical descriptions. I originally collected data on articles in order to test the hypothesis that the order of article and noun correlates with the order of object and verb in that article-noun order correlates

[18] Davis et al are quite correct that it is often difficult to get a precise idea of the meaning of an article from a descriptive grammar; however, since I believe I first examined Kuipers' (1967) description of S̲kw̲x̲wú7mesh about thirty years ago, I was probably aware that the S̲kw̲x̲wú7mesh determiner *ta* does not code definiteness in the standard sense long before Davis, Gillon, or Matthewson were, and I was able to conclude this on the basis of Kuipers' description. In other words, for the particular case that Davis et al focus on, both methodologies uncovered this fact.

with object-verb order and noun-article order correlates with verb-object order, an hypothesis that was suggested by claims that modifiers of nouns tend to precede nouns in OV languages and to follow nouns in VO languages (under the traditional view that articles modify nouns). What I found was the opposite correlation (Dryer 1988, 1992): article-noun order is found more often in VO languages and noun-article order is found more often in OV languages. I would not have been able to reach this conclusion without collecting data from a large number of grammatical descriptions.

8. Methodological issues

Since the primary issues involved in this debate are methodological, there are a couple of general methodological issues raised by Davis et al that bear on the particular issue of definite articles that merit some mention. These are the issue of falsifiable hypotheses and the issue of negative evidence.

While Davis et al consider the issue of falsifiable hypotheses to be a central methodological issue distinguishing their methodology from mine (and others), I do not believe that this is as significant a difference as they seem to believe. For one thing, part of the process of my collecting data on articles can be construed as involving an hypothesis at any given point in time that the range of meanings of articles that I had found so far was the complete range of meanings of articles in any language and each time I discover a language with articles that do not fit this, I am falsifying that hypothesis. Thus both methodologies involve positing falsifiable hypotheses.

Furthermore, when looking at descriptions of individual languages, I frequently look at the characterization of an article given by the author, form an hypothesis as to the meaning and use of that article, and then examine examples and when possible texts to test that hypothesis. More specifically, if an author describes an article as a definite article, then I examine examples to see whether instances of that article correspond to instances of the definite article in the English sentence gloss and, if texts are available, to see whether the use of the article in the sentences in the text is consistent with the use of the definite article in English. Very occasionally, I find something called a definite article used in introducing an entity into a text and if this happens more than once for a language, I decide that it probably is not really a marker of definiteness in the usual sense of this term. Far more commonly, what I find is that something that is called a definite article is frequently not used in contexts where a definite article would be required in English. In such cases, I will decide that the article had a narrower distribution than the definite article in English. In fact, my conclusion that languages with articles that occur in only a subset of the contexts where English uses the definite article are far more common than languages where there is an article with a distribution very similar to the definite article in English is based primarily on inspection of examples, not on explicit statements in grammatical descriptions. It is very common for descriptions to refer to something as a definite article without noting that it has a narrower distribution than the definite article in English.

The one thing I cannot do using my methodology is answer questions that would

require negative evidence, since grammatical descriptions rarely give evidence of that sort bearing on the meaning of articles. So Davis et al are quite accurate in saying that this is a shortcoming of my methodology. For example, there are many languages with articles described as definite articles where all instances in texts provided are anaphoric; but without negative evidence I cannot know whether the article is specifically anaphoric or not. But I have admitted that there are clear limitations on what my methodology can achieve, that it is better suited to providing a bird's eye view of the landscape than precise characterizations of articles in individual languages and one of the sources of those limitations is the unavailability of negative evidence. Furthermore, it is not the case that I have any principled objection to the use of negative evidence when I work on individual languages: some of the arguments in Brown and Dryer (2008), for example, rely on negative evidence.

I do want to point out one risk associated with the particular way Davis et al propose formulating initial hypotheses. The methodology Davis et al advocate in the context of definite and indefinite articles is to initially assume for a language that an article codes definiteness or indefiniteness and then test this hypothesis on the basis of features of definiteness and indefiniteness discussed in the literature. Consider the hypothetical example of applying the tests that Davis et al describe to an article in a language that is restricted to anaphoric definites. As far as I can see, the tests they describe would not falsify the hypothesis that this article coded definiteness. However, Davis et al emphasize that they are only illustrating the sort of tests that they apply and in fact Gillon (2006), in discussing the determiner *ta* in S<u>k</u>w<u>x</u>wú7mesh, does explicitly

address the question whether it can occur in both anaphoric and nonanaphoric definite contexts.[19] But it remains the case that the methodology, as they describe it, will only test hypotheses suggested by distinctions discussed in the literature on definiteness. Consider the more unusual articles described in §5 above. For example, there are a number of languages where there is a distinction between more recently mentioned anaphoric definites and less recently mentioned anaphoric definites. Since this is a distinction that as far as I know has not been discussed in the literature on definiteness, it is not clear how examining such articles in light of the notions discussed in the literature on definiteness would identify this distinction and hence would not uncover the diversity represented by such articles. More generally, if the use of an article in a language is conditioned by some factor not discussed in the literature on definiteness, it is not clear how their methodology will discover what that factor is and since much of the diversity of articles among the world's languages is probably due to factors that are different from those discussed in the literature on definiteness, much of this diversity will remain uncovered by their methodology.

The methodology I describe above that I have used on some languages is to examine texts and see whether a particular article is used in exactly the contexts where the definite article is used in English. This methodology can also be characterized as positing an initial falsifiable hypothesis, and it is a stronger hypothesis than an hypothesis

[19] One approach that is more in the spirit of what Davis et al do would be to develop a fine-grained inventory of uses of the definite article in English (using at least distinctions of the sort made by Hawkins (1978) and Prince (1992)) and then determine for a given article whether it was possible and whether it is obligatory for noun phrases of that sort.

based only on notions discussed in the literature on definiteness. However, once that hypothesis has been falsified (as it has been for most languages I applied this to), the next step is to try to determine exactly what does govern the use of an article. That involves looking for patterns in the data with the goal of coming up with some hypothesis as to what actually govern the use of the article. Crucially, however, these hypotheses will be ones that come out of the data for that language, not hypotheses that one brings to a language based on English or other languages. Unfortunately, because of the fact that grammatical descriptions rarely give negative evidence bearing on the use of article, I am not able to apply this beyond what can be decided on the basis of positive evidence. In short, neither methodology is ideally suited to answering questions about the exact use of particular articles and thus a typology of articles.

9. Conclusion

The primary goal of this response has been to illustrate the extent to which the methodology I have used for my WALS chapters on definite and indefinite articles has uncovered diversity, thus supporting the more general claim that this methodology is well-suited to uncovering at least certain kinds of diversity. For reasons of space, I have not been able to provide evidence for my claims with the degree of detail that I would like, which I plan to do in Dryer (in preparation). The amount of diversity uncovered by my methodology that has not been uncovered by the methodology of Davis et al provides clear evidence against their claim "that hypothesis-driven fieldwork has led to a more complete and accurate picture of linguistic diversity than has been produced by methods

which rely largely on extracting information from existing descriptive grammars". I have conceded that there are aspects of diversity that my methodology is not suited to uncovering and which their methodology is better suited for. My general conclusion is that the optimal way for linguistics, as a discipline, to uncover diversity is to pursue different methodologies, including but not limited to my methodology and the methodology employed by Davis et al.

References

AIKHENVALD, ALEXANDRA. 2003. *A grammar of Tariana, from northwest Amazonia*. Cambridge University Press.

AUSTIN, J. L. 1962. *How to do things with words: the William James Lectures delivered at Harvard University in 1955*, ed. by J. O. Urmson and Marina Sbisà. Oxford University Press

BALLANTYNE, KEIRA GEBBIE. 2005. *Textual structure and discourse prominence in Yapese narrative*. University of Hawaii doctoral dissertation.

BLACKINGS, MAIRI AND NIGEL FABB. 2003. *A Grammar of Ma'di*. Mouton Grammar Library 32. Berlin: Mouton de Gruyter.

BROWN, LEA, AND MATTHEW S. DRYER. 2008. The verbs for 'and' in Walman, a Torricelli language of Papua New Guinea. *Language* 84: 528 - 565.

BRUCE, LES. 1984. *The Alamblak Language of Papua New Guinea (East Sepik)*. Pacific Linguistics: Series C 81. Canberra: Research School of Pacific and Asian Studies, Australian National University.

CARLSON, GREG AND RACHEL SUSSMAN. 2005. Seemingly indefinite definites. In *Linguistic Evidence*, ed. by S. Kepsar and M. Reis, 71-86. Berlin: de Gruyter.

COUPE, ALEXANDER ROBERTSON. 2007. *A Grammar of Mongsen Ao*. Mouton Grammar Library 39. Berlin: Mouton de Gruyter.

DAVIS, PHILIP, AND ROSS SAUNDERS. 1997. *A grammar of Bella Coola*. Missoula, MT: University of Montana Occasional Papers in Linguistics 13.

DAYLEY, JON P. 1985. *Tzutujil grammar*. University of California Publications in Linguistics 107. Berkeley: University of California Press.

DIESSEL, HOLGER. 2013. Pronominal and adnominal demonstratives. *The World Atlas of Language Structures Online*, ed. by Matthew S. Dryer and Martin Haspelmath. Munich: Max Planck Digital Library, Chapter 38. Online: http://wals.info/chapter/38

DRYER, MATTHEW S. 1973. *On the generation of an English description of a picture from a predicate calculus representation*. University of Toronto Masters Thesis.

DRYER, MATTHEW S. 1989. Article-noun order. In *Papers from the 25th Regional Meeting of the Chicago Linguistic Society*, pp. 83-97.

DRYER, MATTHEW S. 1992. The Greenbergian word order correlations. *Language* 68: 81–138.

DRYER, MATTHEW S. 2005a. Definite Articles. In Haspelmath et al (2005), 154-157.

DRYER, MATTHEW S. 2005b. Indefinite Articles. In Haspelmath et al (2005), 158-161.

DRYER, MATTHEW S. 2013a. Definite Articles. *The World Atlas of Language Structures Online*, ed. by Matthew S. Dryer and Martin Haspelmath. Munich: Max Planck Digital Library, Chapter 37. Online: http://wals.info/chapter/37.

DRYER, MATTHEW S. 2013b. Indefinite Articles. *The World Atlas of Language Structures Online*, ed. by Matthew S. Dryer and Martin Haspelmath. Munich: Max Planck Digital Library, chapter 38. Online: http://wals.info/chapter/38

DRYER, MATTHEW S. In preparation. Towards a typology of markers of definiteness, indefiniteness, and related notions.

DRYER, MATTHEW S. AND MARTIN HASPELMATH, eds. 2013. *The World Atlas of Linguistic Structures Online*. Online: http://wals.info/.

ENÇ, MÜRVET. 1991. The semantics of specificity. *Linguistic Inquiry*, 22.1.1-25.

FODOR, JANET AND IVAN SAG. 1982. Referential and quantificational indefinites. *Linguistics and Philosophy* 5. 355–398.

FRANÇOIS, ALEXANDRE. 2002. *Araki: A disappearing language of Vanuatu*. Pacific Linguistics 522. Canberra: Research School of Pacific and Asian Studies, Australian National University.

GERNSBACHER, MORTON ANN, AND SUZANNE SHROYER. 1989. The cataphoric use of indefinite *this* in spoken narratives. *Memory and Cognition* 17: 536-540.

GILLON, CARRIE. 2006. The semantics of determiners: domain restriction in Skwxwú7mesh. University of British Columbia dissertation.

GIVÓN, T. 1978. Definiteness and referentiality. *Universals of Human Language*, ed. by Joseph Greenberg, 291-330. Cambridge: MIT Press.

GIVÓN, T. 1984. *Syntax*: *A functional-typological introduction*, volume 1. Amsterdam: John Benjamins.

GORDON, LYNN. 1986. *Maricopa morphology and syntax*. University of California Publications in Linguistics 108. Berkeley, California: University of California Press.

GUNDEL, JEANETTE K., NANCY HEDBERG AND RON ZACHARSKI. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69: 274-307.

HAGÈGE, CLAUDE. 1981. *Le Comox lhaamen de Colombie Britannique: Présentation d'une langue amerindienne*. Amerindia: Revue d'ethno-linguistique amerindienne, numero special 2, 1981.

HAMEL, PATRICIA J. 1994. *A grammar and lexicon of Loniu, Papua New Guinea*. Pacific Linguistics Series C 103. Canberra: Research School of Pacific and Asian Studies, Australian National University.

HARMS, PHILIP L. 1994. *Epena Pedee syntax*. Studies in the Indigenous Languages of Colombia 4. Summer Institute of Linguistics Publications in Linguistics 118. Arlington: The Summer Institute of Linguistics and the University of Texas at Arlington.

HASPELMATH, MARTIN. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86: 663-687.

HASPELMATH, MARTIN, MATTHEW S. DRYER, DAVID GIL, AND BERNARD COMRIE,

eds. 2005. *The World Atlas of Linguistic Structures*. Oxford University Press.

HAWKINS, JOHN. 1978. *Definiteness and indefiniteness: A study in reference and grammaticality prediction*. London: Croom Helm.

HEUVEL, WILCO VAN DEN. 2006. *Biak: Description of an Austronesian language of Papua*. Vrije Universiteit Amsterdam doctoral dissertation.

HOOPER, ROBIN. 1993. Studies in Tokelauan syntax. University of Auckland doctoral dissertation.

HOOPER, ROBIN. 1996. *Tokelauan*. Languages of the World/Materials 58. Munich: Lincom Europa.

HUALDE, JOSÉ IGNACIO AND ORTIZ DE URBINA, eds. 2003. *A grammar of Basque*. Berlin: Mouton de Gruyter.

KNOWLES, SUSAN MARIE. 1984. *A descriptive grammar of Chontal Maya (San Carlos dialect)*. Tulane University doctoral dissertation.

KUIPERS, AERT. 1967. *The Squamish language: Grammar, texts, dictionary*. The Hague: Mouton.

LOUGHNANE, ROBYN. 2009. A grammar of Oksapmin. University of Melbourne doctoral dissertation.

MATTHEWSON, LISA. 1998. *Determiner systems and quantificational strategies: Evidence from Salish*. The Hague: Holland Academic Graphics.

MUSHIN, ILANA. 2012. *A grammar of (Western) Garrwa*. Pacific Linguistics 637. Berlin: De Gruyter Mouton.

PALMER, BILL. 2009. *Kokota Grammar*. Oceanic Linguistics Special Publication 35. Honolulu: University of Hawaii Press.

PAYNE, DORIS L. 1992. Nonidentifiable information and pragmatic order rules in ’O’odham. In *Pragmatics of word order flexibility*, ed. by Doris L. Payne, 137-166. Amsterdam: John Benjamins.

PET, WILLEM J. A. 1987. *Lokono Dian: the Arawak language of Suriname: A sketch of its grammatical structure and lexicon*. Cornell University doctoral dissertation.

PRINCE, ELLEN. 1981. On the inferencing of indefinite-this NPs. In *Elements of discourse understanding*, ed. by Aravind Joshi, Bonnie Webber and Ivan Sag, 231-50. Cambridge: Cambridge University Press.

PRINCE, ELLEN. 1992. The ZPG letter: Subjects, definiteness, and information-status. In *Discourse description: Diverse linguistic analyses of a fund-raising text*, ed. by William Mann and Sandra Thompson, 295-325. Amsterdam: Benjamins.

ROSS, MALCOLM. 2002. Siar. In *The Oceanic Languages*, ed. by John Lynch, Malcolm Ross and Terry Crowley, 410-425. Richmond: Curzon.

SALTARELLI, MARIO WITH MIREN AZKARATE, DAVID FARWELL, JON ORTIZ DE URBINA, AND LOURDES OÑEDERRA. 1988. *Basque*. Croom Helm Descriptive Grammars. London: Croom Helm.

SAMARIN, WILLIAM J. 1966. *The Gbeya language: grammar, texts and vocabularies*. University of California Publications in Linguistics 44. Berkeley: University of California Press.

SCHROEDER, CHRISTOPHER. 1999. *The Turkish nominal phrase in spoken discourse*. Wiesbaden: Harrassowitz.

SCHUH, RUSSELL G. 1972. *Aspects of Ngizim syntax*. University of California at Los Angeles doctoral dissertation.

SMYE, GRAHAM. 2004. *A grammar of Anufo*. Language monographs 7. Legon [Ghana]: Institute of African Studies, University of Ghana.

WALKER, DALE F. 1976. *A grammar of the Lampung language: the Pesisir dialect of Way Lima*. NUSA 2. Jakarta: Universitas Katolik Indonesia Atma Jaya.

WEDEKIND, CHARLOTTE. 1972. *An outline of the grammar of Busa (Nigeria)*. University of Kiel doctoral dissertation.

WOLFF, EKKEHARDT H. 1983. *A grammar of the Lamang Language (Gwàd Làmàŋ)*. Afrikanistische Forschungen X. Glückstadt: J. J. Augustin.