

## Review Article

**Cite this article:** Kupisch T, Polinsky M (2022). Language history on fast forward: Innovations in heritage languages and diachronic change. *Bilingualism: Language and Cognition* 25, 1–12. <https://doi.org/10.1017/S1366728921000997>

Received: 6 August 2021  
Revised: 2 October 2021  
Accepted: 7 October 2021  
First published online: 8 November 2021

### Keywords:

diachronic change; accelerated change; heritage languages; determiners; (de-) grammaticalization

### Address for correspondence:

Tanja Kupisch,  
E-mail: [tanja.kupisch@uni-konstanz.de](mailto:tanja.kupisch@uni-konstanz.de)

### Abstract

There has been a substantial amount of research on heritage language acquisition and diachronic change. Although recent work has increasingly pointed to parallels between those two areas, it remains unclear how systematic these are. In this paper, we provide a bird's eye view, illustrating how patterns of diachronic change are mirrored in heritage language grammars. In doing so, we focus on one of the best-described grammaticalization processes – namely, the formation of articles from demonstratives and numerals, reviewing studies on heritage varieties which mirror those processes. Based on this review, we make two main points: that change in heritage language can be predicted based on established diachronic scenarios, and that heritage languages often amplify incipient changes in the baseline. After discussing a number of attested changes in a bilingual context, we identify directions for future research in the domain of determiners in heritage languages.

## 1. Introduction

Grammaticalization is a process by which lexical words turn into grammatical markers. For instance, the French noun *pas* “step” has grammaticalized into a negation marker, as in *Je ne sais pas* “I don't know” (see Hopper & Traugott, 1993, 2003 for more examples and general discussion). Grammaticalization can affect the form and function of a linguistic sign. Formally, it may lose some of its phonetic substance (erosion). In functional terms, its original morpho-syntactic properties may change, the original semantic content can bleach out, while its contexts of use are gradually expanded (generalization) as grammaticalizing forms “come to serve a larger and larger range of meaningful morphosyntactic purposes” (Hopper & Traugott, 1993, p. 95). Approaches to grammaticalization have stressed either the process of context expansion of a given linguistic sign (Himmelman, 1997) or the loss of autonomy, accompanied by formal reduction (Givón, 1979; Haspelmath, 2004; Lehmann, 2015). As we will show, the concept of context expansion, along with that of degrammaticalization – that is, the idea that grammaticalization is not unidirectional – is particularly relevant for heritage languages.

We will be concerned with the formation of articles from demonstratives or from numerals: an instance of grammaticalization whereby a closed-class functional item loses some of its semantic content. The evolution of demonstratives into articles constitutes a relatively well-documented example of grammaticalization, with a plethora of evidence from typologically different languages. Heine, Claudi and Hünnemeyer (1991, p. 6) even refer to it as one of the paradigm cases of grammaticalization, because it illustrates the unidirectional development of a demonstrative as it continually expands its range of use until becoming obligatory with a noun, while its original semantic content gradually bleaches out.

While little is known about the rate at which grammatical elements develop, or the factors that influence this rate, we know for sure that the cycle of the definite article evolves at an extremely slow pace. Greenberg (1978, p. 59), for example, mentions the case of Amharic, which in its recorded history of almost 3,000 years has gone through all stages from demonstrative to noun marker. As we will outline in more detail below, the process from the Latin demonstrative to the definite articles that we find in the modern Romance languages has taken at least 1,000 years, and the articles still haven't reached their final destination, although French is arguably close to this final stage, given that it disallows bare nominals almost entirely (see Epstein, 1995; Longobardi, 1999, 2011 for more detailed and critical accounts).

It is hard to document slow changes, especially in the absence of recorded data; this luxury is available for Amharic and Romance, but not for many other languages. In the absence of historical documents, researchers have relied on other means, such as computational modeling of change (see Baker, 2008; Kirby, 2017, for overviews), microvariation across dialects (e.g., Brandner, 2012; Garzonio & Poletto, 2018), or child language data (Lightfoot, 1991, 1999; Yang, 2003, 2016). In this paper, we propose a novel testbed for language change phenomena:

minority or heritage languages.<sup>1</sup> Here, change is accelerated, often occurring from one generation to the next, and can therefore be viewed more vividly than over the course of regular language development, as observed in homeland settings. More specifically, we propose that heritage languages are an ideal testbed for language change (a) because they are in a situation of EXTREME LANGUAGE CONTACT, and (b) because they rely on colloquial input and are typically not subject to standardization, unlike their homeland<sup>2</sup> counterparts. Based on the example of article use, we show that heritage languages undergo the same processes of grammaticalization and degrammaticalization as (other) natural languages do. Therefore, GRAMMATICAL PATTERNS IN HERITAGE LANGUAGES CAN BE PREDICTED ON THE BASIS OF DIACHRONIC CHANGE, and heritage languages can AMPLIFY and foreground developments that are known to take place in language diachrony and are potentially already taking place in the homeland variety (Nagy, 2016, 2017; Polinsky, forthcoming; Rinke & Flores, 2014; Rothman, 2007). By using the term AMPLIFY we imply that such processes can also take place in homeland varieties that are spoken as a national language by the majority of a population, but here they are typically obstructed or even blocked by standardization, so that we do not see significant changes from one generation to the next. In heritage language settings, by contrast, such changes can happen more quickly and thus become visible and quantitatively significant when comparing two generations.

Finally, while grammaticalization studies generally stress universal principles driving language change, it is clear that the stages before European state-building (typically documented in such studies) also involved language contact. In the latter case, however, this could be characterized as ‘normal’ language contact settings, settings of societal multilingualism or cases of ‘external’ language contact. As Classen (2013, p. 131) observes, “people in the European Middle Ages were already extensively on the move” and multilingualism was very common. The difference from today’s heritage language settings is that the latter emerge under conditions of a strong monolingual habitus with a single dominant majority language, which makes them ‘extreme’ compared to the contact settings before and outside of the European nation-state building trend.

Before we proceed, a few general observations about heritage languages and their speakers as well as terminological clarifications are in order. It is widely agreed that a HERITAGE LANGUAGE (HL) is a minority language acquired in a regional or national environment where it represents neither an official language nor the societal language. An HL is acquired in a naturalistic setting, either as the only first language or as one of two or more first languages, through one or several family members. Definitions and descriptions of heritage speakers (HSs) vary (cf. Benmamoun, Montrul & Polinsky, 2013; Fishman, 2001; Flores, Kupisch & Rinke, 2017; Kupisch & Rothman, 2018; Montrul, 2008, 2016; Polinsky, 2018; Polinsky & Scontras, 2020; Rothman, 2009). For the purpose of the present paper, the details of who exactly counts as an HS or not is less crucial, because we expect similar processes

to obtain in situations of EXTREME LANGUAGE CONTACT. We consider language contact to be extreme if an individual acquires two (or more) languages in early childhood, of which one represents a minority language. Given what we know about early bilingual development (e.g., Meisel, 1986, 2011a), minority languages are expected to evolve in a qualitatively similar way as in monolingual children, but the DEGREE OF ISOLATION and the SIZE OF THE COMMUNITY might determine the speed of language development in each individual language. Since the extra-linguistic circumstances under which HLs develop vary substantially, so do, as a consequence, the degrees to which HS grammars differ from each other and from the homeland variety.

Taking diachronic patterns of language change as a starting point, we will show that the same processes occur in HL acquisition but faster. The terminological implication is that what has been labeled ATTRITION OF DIVERGENT ACQUISITION in HLs is no different from INNOVATION OR GRAMMATICALIZATION in the literature on language change. Just as in situations of natural language change, HSs may acquire different target systems from the generation before them because their linguistic experience is different from that of the previous generation. Moreover, we will show that language acquisition in HSs parallels natural language change not only at an abstract conceptual level, but that data from HSs also mirror specific universal patterns of language change. The idea that transfer in language contact situations is regular and follows certain universal patterns is not new and has been advanced both in grammaticalization-oriented work (e.g., Haase & Nau, 1996; Heine & Kuteva, 2005) and by creolists (e.g., Bickerton, 1981). Indeed, Michaelis and Haspelmath (2020) stress that creoles not only show great transparency and simplicity but also accelerated functionalization (content items become lexical items) when compared to their lexifiers (English, French, Spanish, Portuguese). We propose that HS data is another area where a rapprochement of diachronic research and bilingual acquisition research can help us understand how languages change.

In what follows, we will introduce classic proposals on article grammaticalization, providing some well-documented examples of diachronic change. In section 3, we turn to HL acquisition, arguing that innovation in heritage languages mirrors the diachronic patterns. In section 4, we provide an outlook and suggestions for future directions.

## 2. Article grammaticalization

### 2.1. The definite article

The best-known model of article grammaticalization goes back to Greenberg (1978). According to his definite article cycle, the article develops from a demonstrative along a path that includes the following four stages:

The scale correlates with a decrease in referentiality or specificity: an anaphorically used demonstrative in Stage 0 turns into a definite article in Stage 1, which is used to mark a referent as being identifiable (e.g., *The lion I caught yesterday*, or *We spotted a lion. The lion was white*, or *The moon is pale*, or *Close the window*). In Stage 2, this article is extended to specific but unidentified referents (as in English *Mary likes to go to the cinema*, where *the cinema* gets a weak reading with no unique reference), so that the definite also enters the domain of indefinites (Demske, 2020). In this stage, the choice of articles is largely grammaticalized, being determined by syntactic construction (Greenberg, 1978, p. 63). The grammars of languages in Stage 2 habitually list not

<sup>1</sup>Asking whether bilingual language acquisition and language change are linked is not entirely new; see especially Meisel (2011b) and Meisel, Elsig and Rinke (2013). However, the aforementioned work was primarily interested in whether and which types of bilinguals can be the principal agents of diachronic change, while we are concerned with identifying qualitatively similar pathways in the rise and decline of functional categories.

<sup>2</sup>In this paper, we will use the term HOMELAND with reference to the country where the HSs’ parents (or grandparents) were born. This concept does not necessarily coincide with the HSs’ homeland, or even what the immigrants themselves consider to be their homeland.

**Table 1.** Greenberg (1978, p. 61), adapted

Stage 0	Stage 1	Stage 2	Stage 3
Demonstrative	Definite article	Non-generic article	Noun marker

the uses of the article form but rather situations in which it is NOT used (ibid., p. 64). Although Greenberg stresses that we cannot expect all languages with Stage 2 articles to have the same set of uses for the non-articulated forms, some general functions resist the spread of the article: proper names (instances of “automatic” or inherent definiteness), generic uses as in negation and predicates, and generic verb objects (ibid., pp. 64–66). In Stage 3, referentiality no longer plays any role. The article is either a mere gender marker (if the original demonstrative marked gender) or simply a marker of nounhood (ibid., p. 69). This development is often accompanied by a loss of phonetic substance, which can lead to affixation, before the cycle starts anew.

Some aspects of Greenberg’s model have been found controversial, leading to updated versions of the original proposal. The first point of criticism is that the demonstrative turns into a GENDER marker (Greenberg, 1978, p. 55, 62) rather than a marker of DEFINITENESS (see Demske, 2001; Himmelmann, 1997; Leiss, 2000; Lyons, 1999 for other views). The second point under debate has been that the steps from demonstrative via definite article to SPECIFIER are not entirely unproblematic, because in many languages, referentiality and specificity are orthogonal to definiteness (Lyons, 1999). A final point of criticism is that Stage 3 might need to be further differentiated. For example, German definite articles (arguably Stage 2 articles) have some features of noun markers, as they can precede proper names (*die Nora*), which are rigid designators by themselves. At the same time, they also show features of classifiers, as they can be used for socio-pragmatic classifications, such as derogatory uses (*der Trump*) (see Flick, 2021). The details are not crucial, but for our purposes it is important to keep in mind, as pointed out by Greenberg himself, that contexts for bare noun use in Stage 2 may differ across languages.

The evolution of the Latin demonstrative into the definite article in the Modern Romance languages is a well-documented example of definite-article grammaticalization. Classical Latin had no expression of the distinction between known vs. unknown entities. In some cases, the known entity was expressed by *is* or *ille*, and unknown ones by *quidam* or *unus* – that is, postposed *quidam* could optionally signal indefinite specific reference (although no one has ever talked about *quidam* as an article). Data from the 2nd century, such as from *Vetus Latina*, the oldest translation of the Old and New Testaments into Latin, *ille* and *ipse* can only occur with textually known elements, specified in the preceding discourse or understood, but the two are not obligatory in these functions (Renzi, 1976, p. 27).

In the 6th century (Late Latin), such as in *The Rule of Saint Benedict* (a book of instructions for monks dated 516 AD), the article occurs for the first time in contexts where a referent is not identifiable based on text, but by being inalienable (*the hand*), unique (*the moon*) or in the common ground of speaker and hearer (*Feed the cat!*). Renzi (1976, p. 31) sees this as the decisive step for articlehood. In Greenberg’s sense, it could be considered a Stage 2 article at his point. Thus, on the one hand *San Benedetto* bridges Classical and Late Latin, while on the other hand it bridges Late Latin and early Romance, where the article occurs in all of its uses. The article does not occur in either

the French *Strasbourg Oaths* (dated 842) or the Southern Italian *Placiti Cassinesi* (from 960–963). Reasons for its absence might be the brevity of the texts or the writers’ knowledge of Latin. Instead, the article is found, for the first time, in the *Sequence of Saint Eulalia* (from 880), one of the earliest surviving biographical texts written in Old French, and in other ancient Provençal documents from the 11th and 12th centuries. While it may be a matter of debate from what point the former demonstrative can be called an ‘article’,<sup>3</sup> it took more than a thousand years from the onset of functional extension of the demonstrative to the article found in modern Romance languages.

Another well-documented example concerns the German definite article, whose development resembles that of the Romance languages in many ways. The functional change from demonstrative to definite article happened with the initially article-less period of Old High German, conventionally associated with the time from 750 to 1050 (Oubouzar, 1992; Szczepaniak, 2011). In Old High German, definiteness could be expressed by word order, verbal aspect, case, adjective inflection and inherently definite markers, such as possessives and demonstratives. In Early Old High German, the demonstrative *ther* (later *dher* and *der*) was primarily used to single out a referent from a number of potential referents by means of discourse information (anaphoric use) or based on common ground (anamnesic use), as is typical for demonstratives (Himmelmann, 1997, p. 85). Only about 100 years later (9th century), the translator of Tatian<sup>4</sup> makes use of an anaphoric definite, so that a previously introduced referent is mostly taken up by *ther*, as is typical for a Stage 1 article in Greenberg’s sense. But there are additional innovations: identifiable people are now regularly preceded by *ther*, but not animals, plants and inanimate objects (Szczepaniak, 2011, p. 75). In other words, over time, the determiner loses its demonstrative component and acquires the individualizing function. This paves the way for contexts exclusive to definite articles, such as non-referential uses (including generics) and combinations with unique entities, where the article is considered to be an expletive because unique referents are inherently identifiable. The Tatian translator already used definite articles with generic reference, but did not do so consistently. These tendencies continue in the Early Middle Ages. The next step is expansion to the use with proper names, which only takes place in Early New High German, around 1350 (Schmuck & Szczepaniak, 2014; Schmuck, 2020), which has been variable until today. In summary, it took more than a thousand years for the Old High German demonstrative to develop into the definite article of modern (Standard) German, where it corresponds to a Stage 2 article. The use of the article with proper names indicates a very advanced degree of grammaticalization. On the other hand, there are many contexts in which bare nouns are still allowed or required (Longobardi, 1994), even more than in the Romance languages, which speaks to Greenberg’s assumption that Stage 2 articles behave differently across languages.

## 2.2 The indefinite article

In Greenberg’s Stage 2, the definite article turns into a non-generic (specific) article, thereby entering the functional domain

<sup>3</sup>See also Ledgeway (2012, p. 89–96), who argues there that the claim that there are true definite articles in Late Latin is less well supported.

<sup>4</sup>Old High German Tatian is a translation of the Diatessaron (Latin-Old High German), the most prominent early gospel harmony created by Tatian, an early Christian apologist of the Assyrian origin.

**Table 2.** Indefinite article grammaticalization

Stage 0	Stage 1	Stage 2	Stage 3	Stage 4
Numeral “one”	Emergent indefinite	Pragmatic indefinite	Referential indefinite	Extended indefinite

of the indefinite article. Nevertheless, indefinite articles are largely assumed to have their own life cycle. Moreover, the emergence of an indefinite article does not presuppose the existence of a definite one. There are languages that have definite articles but no indefinite ones (e.g., Icelandic, Hebrew, Macedonian), and languages with only indefinite articles (e.g., Kurmanji or Cantonese). Thus, the two articles develop independently, although they may compete or overlap at more advanced stages of grammaticalization when their referential functions are bleaching out (Kupisch & Koops, 2007).

There are fewer studies on indefinite article grammaticalization than on the definite one. The indefinite article is typically derived from the numeral *one*. The numeral expresses an extension of a set, foregrounding that the referent consists of exactly ONE ENTITY, while the article individuates, foregrounding that the noun phrase HAS A SPECIFIC REFERENT. As the English contrast between *one* and *a* shows, some languages have different forms for the numeral and the article. If the numeral and the article have similar forms, they may be distinguished by stress (Turkish) or phonological reduction (Dutch). In Turkish, the difference can further be observed in combinations with adjectives (Kornfilt, 1997, p. 275). The numeral *bir* precedes the adjective (e.g., *bir güzel kız* “one beautiful girl”), while the article *bir* precedes the noun (*güzel bir kız* “a beautiful girl” (Schroeder, 2006, p. 556).

The stages in indefinite-article grammaticalization are presented in Figure 2 (based on Christophersen, 1938; Givón, 1981; Heine, 1997; Hopper, 1987). Stage 0 languages, such as Croatian, Hindi or Finnish, have no indefinite articles. In Stage 1 languages, the numeral “one” is an ‘emergent indefinite’, which may be used to express that a noun phrase is specific, although it is not obligatory in this function. In Stage 2, the indefinite can introduce (new) salient referents into the discourse for further reference, but it is pragmatically restricted in this use (Givón, 1981, p. 36). In Stage 3, the article is used independently of pragmatic functions in all referential contexts. Only in Stage 4 can it be used in non-referential contexts as well, including predicative positions (e.g., *He is a doctor*) and generic uses (e.g., *A lion is a dangerous animal*).

Again, this trajectory is well documented for German. The German indefinite article *ein* exemplifies a Stage 4 article, although until the 8th century *ein* was used exclusively as a numeral (Oubouzar, 1997; Szczepaniak, 2011, 2016). The translator of Isidor has used *ein* only to translate the numeral *unus*. Indefinite referents were translated without *ein* (e.g., *chindk uuir-dit uns chiboran* “child is us born”). In Tatian (early 9th century) there are no indications of a grammatical function of *ein* either, but in some cases *ein* is used to single out an object of a number of similar objects. In this period, the indefinite seems to be restricted to the introduction of a new discourse referent, as is typical for Stage 2. In Otfrid (late 9th century), *ein* was sometimes used with a specifying function, as would be expected of a Stage 3 indefinite article. This includes uses in the plural, which show that the article had already started to bleach out its original quantifying function. Such forms increased in Middle High German, but never entered the Standard language, although existing in some

dialects. Additional evidence for the loss of its numeral function is the use of *ein* before mass nouns (Szczepaniak, 2011, p. 84). However, it is only in Notker (10/11th century) that the indefinite article is used in nonspecific contexts, e.g., in comparisons such as *samo-so in einero uesti* (“like in a fortress”). In order to find generic contexts, which would be diagnostic of a Stage 4 article in Greenberg’s sense, we have to wait until the period of Middle High German (1050–1350).

In Standard German today, plural and mass nouns with an indefinite reading still appear bare. Further still, even today, the grammaticalization of the predicate indefinite articles (with professions), still absent in the 12th century (see 1a), has not been completed. Standard German uses bare nouns before predicate nominals such as (1b), while the addition of an article adds some kind of subjective attitude (e.g., admiration) (1c), except in Southern varieties, where (1c) can be used without such a subjective connotation. Note, however, that in spoken German indefinite articles can cliticize, as in *Max is’n Träumer* “Max is a dreamer”, suggesting that it already has features of a Stage 4 indefinite article.

- (1) a. Dancwart der was marchalc (NL 11,1)  
(Szczepaniak, 2011, p. 85)  
       Dancwart he was groom  
 b. Max ist Professor.  
       Max is professor  
 c. Max ist ein Professor!  
       Max is a professor

In summary, the examples show that indefinite articles have a separate cycle, which resembles that of their definite counterparts so that the gradual extension is from referential to non-referential uses.

### 2.3. Degrammaticalization and construction sites

Typological evidence shows that definite articles typically evolve from demonstratives, while indefinite articles derive from the numeral “one”. The individual steps and the content of each step may differ across languages. A more controversial aspect, alluded to in the introduction, is the (non-)existence of DEGRAMMATICALIZATION (see, e.g., Haspelmath, 2004) – that is, whether languages can take the reverse path of what is illustrated in Figures 1 and 2, moving from affixes to clitics or to free-standing functional elements (grammatical words). Degrammaticalization is much less common than grammaticalization, and some researchers insist on the unidirectionality of change from free-standing grammatical words to affixes, yet the opposite process is documented (see Norde, 2009 for a set of examples ranging from modals to conjunctions). There also seem to be some examples relevant to the present discussion.

With respect to definite articles, Epstein (1995) has pointed out that at a fine-grained level of detail, the grammaticalization cline may not be as unidirectional as it seems to be at first sight. He cites Grevisse (1964, p. 275), who identified cases where the definite article was used in Old French but where it

is no longer used today, such as *faire justice* “do justice”, which in Old French would have been *faire la justice*. While all examples are fixed expressions, which might constitute isolated cases, more instances can be found in the comparison of Old English and Modern English. Epstein (1995, p. 173) cites an example from Mustanoja (1960, p. 253) of a generic NP, where the Old English text contained a definite article while Modern English would use a bare nominal.

- (2) furthest go.PL swa feor norþ swa þa hwælhuntan firrest faraþ  
as far north as the whalers  
“as far north as whalers ever go” (Alfred Oros. 17)

Another intriguing observation concerns the evolution of indefinite articles in German.

In Old High German it was possible to use *ein* before plural and mass nouns (Szczepaniak, 2011). Such instances are particularly interesting with respect to the claim that the original numeral function of *ein* is slowly bleaching out, because plurals imply “more than one” and mass nouns imply non-countability, so the defining semantic feature of the numeral *one* would clash with these nouns. Crucially, in the corresponding modern Standard German NP no article could be used (although in some dialects it can).

In Old High German the definite article could be fused with a preceding preposition (Szczepaniak, 2011, p. 88), which is possible in Modern Standard German but restricted to certain prepositions and depending on register and nominal gender (e.g., *zu der* “to the.DAT.F” > *zur, über die* “over the.ACC.F” > *?übere, hinter die* “behind the.ACC.F” > *\*hintere*). Such fusions were also possible in Old High German but at that time restricted to the preposition *zi* “to”, such as *zi themo* > *zemo* (to the.DAT.M), *zi theru* > *zuru* (to the.DAT.F), *zi then* > *zen* (to the.DAT.PL). In Middle High German these enclitic forms exploded, including with prepositions such as *in* “in”, *an* “on”, *vor* “before”, *über* “over”, *hinter* “behind” and *durch* “through”, in the singular and in the plural. In Early New High German these were still frequent, but the pool of possible clitic clusters decreased over time. Today, regional varieties, such as Swiss German, have retained clusters that are absent from Standard German. In short, some article forms were more likely to fuse in Middle High German than they are today.

A related point is that some languages remain what can be thought of as GRAMMATICALIZATION CONSTRUCTION SITES for an extended period of time. Such construction sites can be identified by divergent judgments intra- and inter-individually and brought to light by means of processing studies. A case in point is the use of definite articles with plural subject NPs in German. While most Germans prefer the sentence (3a) under a generic reading where cats are generally intelligent, some speakers also accept (3b) as generic, although the same sentence can also have a specific reading.

- (3) a. Katzen sind intelligent. (generic)  
cats.PL are intelligent  
b. Die Katzen sind intelligent. (generic/specific)  
the.PL cats.PL are intelligent.

Examples of definite article use with generic subject nominals in German, similar to (3b), have been provided by a number of scholars (e.g., Dayal, 2004; Krifka, Pelletier, Carlson, ter Meulen, Link & Chierchia, 1995). An empirical study across different regions in Germany confirmed that the definite determiner is accepted with generic plurals, although bare nominals are by far preferred

(Barton, Kolb & Kupisch, 2015). However, Czymionka and Kupisch (2019), while reaffirming that German speakers are more inclined to interpret sentence like (3b) as specific, showed interesting results for reaction times. Compared to bare nominals and demonstrative-modified NPs, NPs with definite articles took longer to process, which may reflect an ongoing change in the grammar of German. Since the definite plural article has become ambiguous, its processing is more costly.

### 3. Article innovation in heritage languages

We have shown how definite and indefinite articles evolve from demonstratives and numerals, respectively; these processes can stagnate or, in some rare cases, even go in the reverse direction. In this section, we will link previous findings on diachronic change to developments in heritage languages.

#### 3.1. Indefinite article innovation

The following example from Molise Slavic illustrates accelerated language change in a situation of absolute language contact – that is, the situation of a minority language whose speakers are all bi- or trilingual. Molise Slavic, which has evolved as an isolated minority language over a few hundred years, is a variety of Serbo-Croatian traditionally spoken mostly in three villages in Molise, a region in South-Central Italy (Breu, 2012). Today, there are about 1,700 speakers, whose ancestors came from Dalmatia about 500 years ago. Since then, Molise Slavic has been under the influence of Molisian (an Italian dialect), later joined by Standard Italian about 150 years ago after the unification of Italy in 1861. Molise Slavic is mainly used orally, Italian has been the only dominant high variety (‘Dachsprache’), and the closest related language, Standard Croatian, is unintelligible to untrained speakers.

As the examples below illustrate, Molise Slavic uses an article in all contexts that cover the crucial stages in the life cycle of an indefinite.<sup>5,6</sup>

- (4) Sfe skup je uliza na ljud tusti.  
all together AUX enter.PFV.PTCP.SG.M NA person fat  
“Suddenly a fat man came in.” (presentational, Stage 2)
- (5) Ja jiskam na mičicij.  
1SG search.1SG.PRS NA friend  
“I’m looking for a friend.” (a specific one) (specific indefinite, Stage 3)
- (6) Ja ču jimat na mičicij.  
1SG want have.INF NA friend  
“I would like to have a (any) friend.” (nonspecific indefinite, Stage 4)
- (7) Na Zlav ne goriva laž.  
NA Slav not say.3SG.PRS lies  
“A (real) Molisian Slav does not lie.” (generic, Stage 4)

All of this would not be too surprising, if it weren’t the case that within the span of about 500 years Molise Slavic has evolved from

<sup>5</sup>Examples are from Breu (2012). In the glosses, we represent the determiner just as NA.

<sup>6</sup>While not directly related to the focus of this paper, note that Molise Slavic has post-nominal modifiers, which also attests to strong language contact (in Slavic languages, modifiers are typically prenominal).

Serbo-Croatian, an article-less language. In most Slavic languages, including Serbo-Croatian, article functions can be expressed by demonstratives and numerals, and the rare instances of true articles are explained by language contact (Breu, 2012). Interestingly, Breu (2012, p. 301) mentions the possibility of using a numeral in Serbo-Croatian in order to oppose an indefinite referential reading from a referential reading, e.g., *I am looking for a* vs. *the [=my] boyfriend*. However, the numeral has not become obligatory in this function. By contrast, Molise Slavic has developed an indefinite article, which has the exact same functions as its counterpart in Italian.

Another example where language contact seems to have accelerated the expansion of the indefinite articles comes from American Norwegian (AmNo), a heritage language spoken in the United States and Canada (Kinn, 2020). Homeland Norwegian (European Norwegian) allows bare, singular nouns in some contexts where English does not. The most striking difference concerns post-copular, singular predicate nouns: in English, most such nouns must appear with an indefinite article, while European Norwegian uses bare nouns when the predicate is, for example, a profession, role or nationality.

- (8) a. Han er lærer (European Norwegian)  
 b. Han er en lærer (AmNo)  
 he is (a) teacher  
 "He is a teacher."

Despite intense contact with English, which uses the indefinite article, most American Norwegian speakers have retained bare nouns, the pattern of Norwegian as spoken in Norway. However, a minority of the speakers use the indefinite article to some extent. Kinn interprets this as cross-linguistic influence or attrition (a change during the lifetime of individuals rather than divergent attainment causing systematic, parametric change in the Norwegian grammar of these speakers). However, such a change can also take place either without contact or under contact with a majority language that does not use articles in the equivalent construction. For example, Bavarian uses the pattern in (9a), even though the major contact language, Standard German, uses (9b); see also (1).

- (9) a. Der Gert is a Lehrer und koa Fischer  
 the Gert is a teacher and no fisherman  
 b. Gert ist Lehrer, nicht Fischer.  
 Gert is teacher not fisherman  
 "Gert is a teacher, not a fisherman."

### 3.2. Definite article innovation

Incipient stages of definite article use can be illustrated by heritage Mandarin Chinese in the Netherlands, which has extended its use of demonstratives to contexts in which they would be absent in the variety spoken in China (Aalberse, Zou & Andringa, 2017).

Mandarin Chinese has no dedicated morphology to encode definiteness, although definiteness can be expressed via word order and context, and the type of verb plays an additional role (Sybesma, 1992, pp. 176–178). For example, post-verbal bare NPs can in principle receive an indefinite, a definite or a generic reading, but if the verb expresses an unbounded state (e.g., *like*), bare nouns are interpreted as generic. Preverbal bare nouns can only be interpreted as generic or definite, but never as indefinite. Definiteness can also be expressed by the use of possessives,

demonstratives or a demonstrative plus a classifier, in both preverbal and postverbal position, but demonstratives do not have definiteness marking as their primary function.

Dutch has dedicated morphology to encode definiteness. When a referent has been mentioned in the previous discourse or is commonly known, the definite article is obligatory (*het/de* in the singular, *de* in the plural). A new singular referent which is unique in the discourse needs the indefinite *een*. Demonstratives also exist in Dutch but they have a more specialized meaning. In short, Dutch and Mandarin both encode definiteness, but they have different means for doing so. Since both languages can mark definiteness using demonstratives, a possible scenario when these languages are in contact is an overuse of demonstratives in heritage Mandarin Chinese as a result of functional extension.

Aalberse et al. (2017) recruited 12 Mandarin Chinese-speaking families living in the Netherlands (first and second generation) for a story narration task. One parent (first generation) and one child from (second generation) from each family were included in the analysis. The first-generation speakers were native speakers of Mandarin born in mainland China or Malaysia who had moved to the Netherlands for study or work and have lived in the Netherlands ever since. The second-generation speakers were either born in the Netherlands or moved to the Netherlands before age 7. The study also included a control group of seven Mandarin homeland speakers living in mainland China. Thus, the authors compared first-generation Mandarin Chinese speakers, second-generation HSs and a control group in the homeland. The analysis of the narratives showed that all groups used demonstratives more often when a referent was mentioned for the second time than when it was mentioned for the first time. However, with respect to a previously-mentioned referent, there were significant differences across groups. Both groups in the Netherlands used more demonstratives than speakers from China, and amongst the two groups from the Netherlands, the second-generation speakers used still more demonstratives than the first generation (*ibid.*, p. 38).

The example shows that Dutch–Mandarin language contact has led to increased use of demonstratives in Mandarin Chinese in the Netherlands. While it is likely that the presence of obligatory articles in the dominant language Dutch has catalyzed this change, such changes may be expected to occur even in the absence of language contact, albeit slower. In fact, Chen (2004) suggests that the demonstratives in Mandarin Chinese are increasingly used in non-deictic contexts.

For our next example, one of the best documented examples of article use in HSs, we return to the phenomenon introduced in (3) – namely, generic subject DPs. As illustrated in (9a,b), English and German tend to use bare nominals generic subjects in the plural, while French, Italian and Spanish require a definite article, as illustrated by the contrast between (9c–e) and (10c–e).

- (9) a. En. Cats are intelligent. (generic)  
 b. Ge. Katzen sind intelligent. (generic)  
 c. Fr. \*Chats sont intelligents.  
 d. It. \*Gatti sono intelligenti.  
 e. Sp. \*Gatos son inteligentes.
- (10) a. En. The cats are intelligent. (specific)  
 b. Ge. Die Katzen sind intelligent. (specific/?generic)  
 c. Fr. Les chats sont intelligents. (specific/generic)  
 d. It. I gatti sono intelligenti. (specific/generic)  
 e. Sp. Los gatos son inteligentes. (specific/generic)

The obligatory use of the definite article in Romance implies that it is ambiguous between a specific and a generic reading, while the two Germanic languages can distinguish these readings by the presence or absence of the article, although, as discussed above, there are indications that German is moving towards the Romance configuration; see (9b) and (10b).

In a number of independent studies, Serratrice, Sorace, Filiaci and Baldo (2009), Montrul and Ionin (2012), Kupisch (2012), and Barton (2016) examined preferential article use when the Romance language is acquired as a minority language. Regardless of methods and populations, all the studies found that HSs were more inclined to interpret definite subjects as specific and more willing to accept bare nominals, as in (9c-e), as grammatical when compared to a control group. These findings show that Germanic-Romance language contact has led to an increased acceptance of bare nominals in generic subject DPs in French and Italian in Germany, Italian in the UK, and Spanish in the US. Again, it is likely that the absence of obligatory articles in the dominant language has catalyzed this change. Again, the question is whether the change could have occurred in a majority language setting. The English example for degrammaticalization in (2) would suggest that this is indeed possible.

One could argue that in all of the cases we have illustrated so far, it is impossible to tease apart cross-linguistic influence from language-internal or universal mechanisms that have been driving the processes at hand. What if all the cases we have presented here were triggered exclusively by cross-linguistic influence? The aforementioned study by Serratrice et al. (2009) is relevant in this respect. It took as its starting point the observation that English-dominant Italian bilinguals differed from monolingual Italian speakers in overaccepting ungrammatical bare nouns – possibly due to influence from English, which allows bare nouns in these contexts. However, Serratrice and colleagues also tested Italian-Spanish bilinguals on the same property, Spanish and Italian being similar in the distribution and semantics of articles; see (9d,e) and (10d,e). Only the (younger) bilingual children showed visible differences from monolinguals in accepting ungrammatical bare nouns, but these differences did not reach significance. Thus, the comparison between the two groups of bilinguals indicates that cross-linguistic influence is not the only factor at play, although the trend suggests that this kind of comparison is worth investigating further. From a conceptual point of view, an exclusive impact of cross-linguistic influence is not plausible either, since developments always have to work within the grammatical conditions of a language.

We would like to close this section with an illustration that HLs can follow a principled course even when contact effects are ruled out; this is a clear case of accelerated grammaticalization. Norwegian double-definite constructions are a case in point (Van Baal, 2020). In these constructions, European (homeland) Norwegian uses a pre-adjectival, free-standing article and a definite suffix, while American Norwegian speakers replace the pre-adjectival article with a demonstrative, even in contexts where a deictic reading would be excluded.

- (11) a. **den** hvite hest-**en** (European Norwegian)  
the white horse-DEF  
“the white horse”  
b. **denne** hvite hest-**en** (American Norwegian)  
this white horse-DEF  
“the white horse”

The prenominal demonstrative is perceptually more salient than the prenominal article; given the choice, the use of more salient forms is characteristic of HLs more generally (Polinsky, 2018). Arguably, the use of the demonstrative could indicate renewal in the grammaticalization cycle. Crucially, the influence from English, the majority language of these speakers, is unlikely to have caused this change (English speakers would not use demonstratives in such contexts, unless they wanted to express deixis or a special connotation). Further still, this may be yet another case where a heritage language amplifies tendencies incipient in the baseline; demonstrative forms are also sometimes found in homeland dialects from relevant areas. All told, this is an illustration of change that is free of contact effects.

## 4. Discussion

### 4.1. Predicting change without a long view

So far, we have highlighted that data from HSs constitute a magnifying glass through which we can view language change. The reason why it is plausible for the two settings to be similar is because both are a subtype of natural languages (Kupisch & Rothman, 2018; Polinsky, 2018, forthcoming; Rinke & Flores, 2014). Along these lines, it has been proposed that since HSs rely on colloquial input sources, their language may indicate ongoing diachronic change that is hard to identify in the standard variety of a language (Flores & Rinke, 2021; Pires & Rothman, 2009; Rothman, 2007). The reason might be that homeland speakers are subject to a constant standardization process, which ensures, such as through educational institutions or linguistic ‘academies’ (e.g., *Accademia della Crusca* in Italy, *Académie Française* in France) that some kind of agreed-upon norm is maintained. The fact that HLs may amplify trends that are hard to detect in the homeland language makes them an ideal testbed for micro-comparative language investigations, where one or more heritage varieties is compared to the baseline language in the diaspora (the language of first-generation immigrants) and the ‘source’ language in the homeland of their ancestors (e.g., Nagy, 2016, 2017).<sup>7</sup>

Conversely, divergences between an HL and baseline can be predicted on the basis of well-known language change scenarios. The challenge is to anticipate patterns of change and to rule out other patterns. In that regard, Greenberg’s model of grammaticalization provides us with a useful guide for predicting change, and the present paper was a first attempt at fleshing out this kind of predictive approach.

The examples of language change covered here also illustrate a trade-off in complexity, an important point, which challenges the much-too-common view of language under contact as always resulting in simplification. For example, a comparison between Latin and Early Romance with respect to their determiner systems uncovers the following generalizations. Latin had fewer determiner forms (less complexity in the inventory of lexical items), but this went along with more ambiguity and thus more complexity in interpretation, as each form had multiple meanings. The Early Romance determiner systems had developed articles, hence more complexity in surface exponence and the overall DP-structure, but for each article retrieval of the intended referents was more straightforward, hence there was less complexity

<sup>7</sup>In some instances, homeland varieties do not exist; endangered indigenous languages, whose speakers are ‘immigrants in their own land’, are a case in point (see Polinsky, 2018, for discussion).

in interpretation. The Modern Romance systems, where, for example, definite articles have become ambiguous between specific and nonspecific reference, and may eventually become affixal noun markers, are moving back to formal homogeneity (less complexity) but greater complexity in interpretation. Thus, the decrease in the complexity of form is accompanied by greater complexity in interpretation. That entails that no diachronic stage in the development of article systems is more complex or simpler than another. Similarly, no heritage language is more complex or simpler than the baseline associated with it; again, it is a trade-off between complexity of form and complexity of interpretation.<sup>8</sup>

The observation that HL data mirror or amplify ongoing changes in the baseline language has been made previously with various degrees of explicitness by a number of researchers. The genitive of negation in modern spoken Russian is an illustrative example. This phenomenon, albeit popular among linguists (see Harves, 2013 for an overview and further references), is not prominent in the modern baseline (Comrie, Stone & Polinsky, 1996, pp. 146–147), and the heritage varieties barely use the genitive of negation. Inspecting particular patterns of divergence between a HL and its baseline allows us to pinpoint incipient changes in the latter. Conversely, by looking at what is unstable in the baseline we can make predictions about the specific changes that can occur in contact varieties, HLs in particular.

With this latter idea as our guiding principle, can we make more predictions with respect to determiner systems? Based on the diachronic data we have, we raise a number of questions for future research.

#### 4.2. The fate of determiners: some future directions

The first question we would like to raise has to do with the pace of article development: DO DEFINITE AND INDEFINITE DETERMINERS DEVELOP AT THE SAME TIME? Based on diachronic scenarios, it appears that definite determiners are more ‘aggressive’ in change. The prediction that definites are more likely to evolve can easily be tested in language contact settings where the minority language has no articles (and no incipient grammaticalization tendencies are evident), while the majority language has fully developed definite and indefinite articles. In our discussion of Mandarin Chinese spoken in the Netherlands, we noted that there is a growing use of demonstratives with noun phrases whose referent has been introduced in discourse. We can also expect that Mandarin Chinese speakers dominant in Dutch (or any other language with a full-fledged article system) may extend the use of the numeral ‘one’ beyond contexts where simple counting is needed, thus attributing to it some functions of an indefinite article. To reiterate, historical examples of indefinite articles developing from the numeral ‘one’ abound, but in the context of language contact, such a development may be accelerated.

Even if we concentrate on definite articles alone, the range of their functions varies significantly across languages, and that range of functions may also be implicated in language change. On the one end of the spectrum, we find languages like Italian and Portuguese, which use definite articles with proper names,

<sup>8</sup>We are aware that the definition of complexity is far from straightforward. Another example is the diachronic development of future tense forms in French and other Romance languages: from Latin *amabo* via Vulgar Latin *amare habeo* and French *j’aimerai* to contemporary Colloquial French *je vais aimer*. Arguably, analytic constructions are formally more complex as they consist of more items, but they may be easier to parse thanks to more straightforward mapping between form and meaning. The preponderance of analytic constructions in HLs and creole languages deserves more attention.

before possessive pronouns, and with generic plurals; on the other, there are languages like Mangarayi where definite articles carry the typical conservative functions of anaphoric reference (Merlan, 1982). Assuming the universal nature of cognitive statuses involved in the use of referring expressions (Gundel, Hedberg & Zacharski, 1993), it is conceivable that language contact could lead to an expansion of the functions of the definite determiner, from the canonical marker of a uniquely identifiable anaphoric referent to a generalized marker of a free-standing noun phrase. At present, we lack comprehensive empirical data on the development of these functions under contact in general and in HLs in particular.

The mention of Mandarin Chinese above also compels us to consider the distinction between LANGUAGES WITH AND WITHOUT CLASSIFIERS. Mandarin Chinese, where a numeral cannot occur without a classifier, as shown below, is a commonly-cited example of the former:

- (12) liǎng \*(běn) shū  
two CLF book  
“two books”

In our discussion, we are primarily interested in sortal classifiers (as opposed to mensural classifiers). The latter are akin to collective nouns, such as *pound* or *drop*, that encode inherent characteristics of an object, such as shape, size or function, and their main function is to atomize the referent of a given noun – that is, to make it countable.

Some researchers explain the difference between classifier languages like Mandarin Chinese and non-classifier languages such as Romance or Germanic by differences in the nominal system, arguing that only non-classifier languages have a mass/count distinction among nouns, while classifier languages do not (Chierchia, 1998; Wilhelm, 2008). Other researchers attribute the difference to the content and structure of the numerical expressions (Krifka, 1995). It is not our goal to adjudicate between the two explanations (but see Bale & Coon, 2014); rather, we are interested in the interaction between classifiers and demonstratives. Despite differences in the structure of classifier phrases across languages (e.g., Cheng & Sybesma, 1999; Saito, Lin & Murasugi, 2008), the noun phrase in classifier languages is embedded below a classifier phrase, and classifiers can perform the deictic function, one that is assumed to be carried out by determiners in Romance and Germanic languages. While classifiers can or even must co-occur with demonstratives (see Cheng & Sybesma, 1999, p. 530 on Mandarin; Jenks, 2011 on Thai), the question is whether they could co-occur with articles of the Romance/Germanic type and whether this would eventually lead to the disappearance of classifiers (presumably because of redundancy in the marking of deixis). The literature on classifier use in HLs, with English as the dominant language, suggests that the inventory of classifiers shrinks but the category does not disappear, as has been demonstrated for Cantonese (Nagy & Lo, 2019; Wei & Lee, 2001). Rather, there is a trend toward the over-generalization of a generic classifier, more commonly used with singular nouns (Nagy & Lo, 2019). Assuming that under contact with a language that has articles, demonstratives gain importance as the exponent of definiteness (and numerals arise as exponents of indefiniteness), it is possible that composite expressions expressing (in)definiteness will arise, thus:

- (13) a. DEM+CLF<sub>[definite]</sub> noun  
b. NUM+CLF<sub>[indefinite]</sub> noun

We would like to underscore that so far, we have only considered cases where the contact language has articles. It is less clear what may happen in a classifier HL when in contact with another language that has classifiers (and no articles), such as heritage Chinese in Vietnam or in Korea where the dominant language also employs classifiers.

### 4.3 The dominant language

Finally, we would like to turn to an important point that often gets overlooked in bilingual context: THE KNOWLEDGE AND USE OF THE DOMINANT LANGUAGE BY HSs. Most studies of the relevant dominant language in a bilingual dyad are based on investigations of monolingual grammars, and it is often tacitly assumed that a bilingual/HS dominant in that language does not deviate from that monolingual grammar. Sound systems and their categorization have long been an exception to such an assumption, as bilinguals have been shown to differ from monolinguals in both languages (see Chang, 2021; Polinsky, 2018, for overviews). There is mounting evidence that a bilingual's divergence from monolingual norms may begin to occur early in bilingual development and, furthermore, persist despite weak proficiency in and infrequent use of the other language. This has been shown repeatedly in studies on child bilingualism. For example, Kupisch (2007) has shown that the acquisition of articles in German (dominant majority language) is accelerated under the influence of Italian (HL). Another special case in point is the new variety of German called *Kiezdeutsch* (Wiese, 2013). This variety has emerged in a special multi-ethnic context in Germany, amongst young speakers in larger urban areas, who typically show a particularly high tolerance of linguistic variation. *Kiezdeutsch* reflects the internal tendencies of German, pointing to an acceleration of internal language change and making it a 'pioneer dialect.' Further, comparing to language use in a more monolingually German neighborhood, *Kiezdeutsch* shows a tendency towards more nonstandard patterns (Wiese & Rehbein, 2015). In short, the assumption that HSs do not differ from monolingual native speakers in their dominant language has become more questionable, thus increasing the impetus to directly test both languages in HSs.

In the domain of referential expressions, it has been observed, perhaps somewhat surprisingly, that while monolingual English speakers allow scope ambiguities in clauses with two quantified expressions, such as (14), English-dominant HSs with different home languages allow only the surface scope, such as (14a); see Ronai (2018) and Scontras, Polinsky, Tsai and Mai (2017):

- (14) A player hit every soccer ball
- There was a single player who hit every soccer ball (surface scope,  $\exists > \forall$ )
  - For every soccer ball, there was a separate player who hit it (inverse scope,  $\forall > \exists$ )

Since the testing of competence in the dominant language by HSs is in its infancy, the possibilities are limitless. Generic subjects with and without articles, which separate the Romance and Germanic languages in a well-known divide (see examples (9) and (10) above) are a prime area of study. Another possible domain has to do with the interaction between negation and indefinite articles, as in (15a, b), where bilinguals may differ from monolinguals in allowing only surface scope (note that scope is also relativized to the structural position of the indefinite noun phrase):

- (15) a. I did not see a security guard (NEG >  $\exists$ ;  $\exists$  > NEG)  
 b. A security guard never stands in this location ( $\exists$  > NEG, NEG >  $\exists$ )

If the quantifier scope studies cited above are any indication, we would expect English-dominant HSs to maintain surface-scope readings only, a prediction that could be tested in multiple language dyads.

The interaction between articles, number marking and focus-sensitive operators is another promising testbed for the status of articles in the dominant language. Consider the following four items:

- (16) a. %Only hyenas hunt at night  
 b. Only the hyenas hunt at night  
 c. Only the hyena hunts at night  
 d. \*Only hyena hunts at night

While (16d) is not accepted by English monolinguals and may not need to be considered, the contrast between the remaining three expressions deserves to be explored; would HSs reject or accept (16a)? Von Stechow (1997) argues that *only* in (16a) is not a determiner, but simply a focus-sensitive operator applied to a bare plural; some English speakers reject that without a context that supports discussing this kind of bare generic claim. If heritage speakers are less sensitive to context, they may be more accepting of (16a). Further still, the acceptance or rejection of (16a) may be dependent on the status of articles in the minority language; if an English-dominant speaker of Norwegian treats (16a) differently from an English-dominant speaker of Spanish (which does not allow bare plural subjects) or Korean (which does not have articles at all), that would be significant evidence in support of bidirectional influence between the languages in a bilingual dyad. Next, in the choice between (16b) and (16c), would HSs prefer one over the other? This latter issue brings us to the next point: the interaction between determiners and agreement categories associated with nouns – namely, number and gender. We take this up in the next subsection.

### 4.4. Determiners do not live alone

We have so far applied broad strokes to show the importance of Hs for our understanding of diachronic change. But determiner marking is often paired with number marking (as in (16a–d)) or gender marking. To illustrate the interplay of all three categories – (in)definiteness, gender and number – consider the diachronic change in Swedish, where strong neuter nouns in the definite plural form ended in *-na*, e.g., *äpple-na* “the apples”. Subsequently such neuter definite plurals were reanalyzed as ending in *-a* (*äpplen-a*), where the form without *-a* (*äpplen*) is interpreted as indefinite (Norde, 2009, p. 65 and references therein).

The interplay of (in)definiteness and number suggests that on close scrutiny, one could expect a more pronounced rise or fall of determiners in the singular, or in one of the genders found in a particular language. For example, it is known that determiners spread from singular to plural contexts, arguably because singulars are more easily construed as unique and identifiable, thus inherently encoding the core features of definite articles. Similarly, the indefinite article starts out denoting “one”, which makes it more easily compatible with singulars. As for gender, one could expect that in a language that has a separate gender category for mass nouns, this category will be more resistant to determiner spread compared to other gender categories. Once again, such uneven development may be more visible in an HL

than in the baseline. As for reverse grammaticalization, if a language has a gender class for mass nouns and uses articles productively in this class, as do some central Italian varieties (e.g., Franco, Manzini & Savoia, 2015), we could expect bare nouns to expand faster in this category, and this trend could become more noticeable under the influence of a contact language in which mass nouns tend to be bare, such as German or English.

#### 4.5. The (universal) nature of change

We proposed that specific structural properties of HLs reflect patterns of diachronic change and that some changes may be reversed (degrammaticalization). But are all developmental mechanisms of the same nature? If change means that existing constructions increase or decrease in their frequency of use, this refers to a predominantly quantitative change, although it might indicate enhanced processing facility. If, on the other hand, we are dealing with grammaticalization patterns as the ones discussed in section 2, we are looking at developments that do reflect changes in the underlying grammatical systems. What is the type of underlying logic determining particular patterns of change? Language contact plays an important role, as we know from the endless discussion of cross-linguistic influence in HSs and bilingualism more generally. But fundamental alterations of grammars cannot be triggered by contact exclusively, because some patterns of change are unrelated to the properties of the contact language, as shown by the American Norwegian case illustrated in (11). Thus, besides language contact, these mechanisms are likely to include language-specific grammatical properties, because some domains of grammar are more susceptible to alterations than others, and this may result in a kind of instability, as suggested in the long debate on the syntax–discourse interface.

The idea that certain language-acquisition settings reveal a time-lapse picture of diachronic developments has previously been discussed in creole studies (cf. Lefebvre, White & Jourdan, 2006 for a discussion and a number of contrasting views). This raises the question to what extent the situation of HLs is comparable to that of creole languages – a question that, as far as we know, has not been addressed. For example, the idea that developmental patterns attested in lexifier languages are responsible for the emergence of properties of creoles has been supported by many creolists. As for determiners, creoles exhibit striking similarities across different languages, independently of which European language they are derived from, the typical pattern being a definite article for presupposed-specific NPs, an indefinite article for asserted-specific NPs, and zero for nonspecific NPs (Bickerton, 1981, p. 5). What is relevant for explanations of diachronic patterns and degrammaticalization is the fact that pidgins typically lack determiners altogether. Therefore, for creoles developing out of pidgins, the pidgin cannot be the source of the creole article/determiner system. Thus, a comparison with creole formation could help us reveal the nature of the change. Bryun (2009) shows that Sranan (Surinamese) definite articles underwent an ‘ordinary’ process of grammaticalization, while the indefinite article grammaticalized rather abruptly in comparison to similar processes in languages with an ordinary history, suggesting that the speed of the change might be an indicator of the underlying process that can help us tease apart contact-induced change (associated with accelerated speed) from language internal change (associated with lower speed) that is driven by universal tendencies that can potentially happen in the absence of language contact.

## 5. Conclusions

We have outlined parallels between well-described diachronic patterns within the domain of nominal morphology and heritage language acquisition, proposing that HL data provides us with opportunities to zoom into language change scenarios that normally take place over a very long period of time. In particular, HL data can foreshadow diachronic change in monolingual settings. Conversely, models of diachronic change can be used to make predictions about HL acquisition. They can further help us understand the role of the contact language in HL acquisition, which is present to different degrees in diachronic settings.

We have used the rise and fall of determiners to illustrate our main points. While a great deal is known about cycles in the diachrony of definite and indefinite determiners, their status in HLs has not been investigated. One particularly interesting point of comparison concerns potential cases of degrammaticalization. HL studies have often documented scenarios where the homeland variety has an overt marker that is being lost in the HL, as with the definite plural article in generic NPs in the Romance language when in contact with a Germanic language. In the context of diachronic change, cases of degrammaticalization have been more controversial (e.g., Fischer, Norde & Perridon, 2004) because they go against the unidirectionality generally observed in classic grammaticalization studies (e.g., Hopper & Traugott, 1993). However, there is mounting evidence that such cases are more common than originally assumed (see, e.g., Norde, 2009). We hope that our work will stimulate explorations in this new area. In our discussion we have outlined possible directions for future research that have so far been largely unexplored.

HLs are a steady presence in the global setting, but in reality most work on HLs has been carried out in the context of immigration to Europe and North America, where the dominant languages (German, French, Spanish, English) have ubiquitous definite and indefinite determiners. At the same time, the majority of immigrants to Europe and North America speak languages without articles (Slavic languages, Chinese, Hindi, Korean, Sinhala, Berber languages), with only one article (Tagalog, Hebrew), with definite affixes (Kurdish, Somali) and with classifiers (Chinese, Korean, Vietnamese, Thai), leaving us with a plethora of opportunities to test those trajectories that the known diachronic patterns would lead us to expect.

**Acknowledgments.** This paper was first presented as a talk and commentary at the RUEG conference Dynamics of Language Contact in February 2021. We wish to thank Ad Backus, Walter Breu, Martin Haspelmath, Tania Ionin, Kari Kinn, Jaklin Kornfilt, Jürgen Meisel, Christoph Schroeder, Greg Scontras, George Walkden, and Heike Wiese for commenting on an earlier draft of this paper.

## References

- Aalberse S, Zou Y and Andringa S** (2017) Extended use of demonstrative pronouns in two generations of Mandarin Chinese speakers in the Netherlands. Evidence of convergence? In Blom E, Cornips L and Schaeffer J (eds), *Crosslinguistic Influence in Bilingualism: In Honor of Aafke Hulk*. Amsterdam: John Benjamins, pp. 25–48.
- Baker A** (2008) Computational approaches to the study of language change. *Language and Linguistics Compass* 2, 289–307.
- Bale A and Coon J** (2014) Classifiers are for numerals, not for nouns: Consequences for the mass/ count distinction. *Linguistic Inquiry* 45, 695–707.
- Barton D** (2016) *Generische Nominalphrasen bei deutsch-französischer Zweisprachigkeit*. Unpublished PhD dissertation, University of Hamburg.

- Barton D, Kolb N and Kupisch T** (2015) Definite article use with generic reference in German: an empirical study. *Zeitschrift für Sprachwissenschaft* **34**, 147–173.
- Benmamoun E, Montrul S and Polinsky M** (2013) Heritage Languages and Their Speakers: Opportunities and Challenges for Linguistics. *Theoretical Linguistics* **39**, 129–181.
- Bickerton D** (1981) *Roots of language*. Ann Arbor MI: Karoma.
- Brandner E** (2012) Syntactic microvariation. *Language and Linguistics Compass* **6**, 113–130.
- Breu W** (2012) The grammaticalization of an indefinite article in Slavic Micro Languages. In Wiemer B, Wälchli B and Hansen B (eds), *Grammatical replication and grammatical borrowing in language contact*. Berlin: De Gruyter, pp. 275–322.
- Bryun A** (2009) Grammaticalization in creoles. Ordinary and not-so ordinary cases. *Studies in Language* **33**, 312–337.
- Chang CB** (2021) Phonetics and phonology. In Montrul S and Polinsky M (eds), *The Cambridge handbook of heritage languages and linguistics*. Cambridge: Cambridge University Press, pp. 581–612.
- Chen P** (2004) Identifiability and definiteness in Chinese. *Linguistics* **42**, 1129–1184.
- Cheng L and Sybesma R** (1999) Bare and not-so-bare noun and the structure of NP. *Linguistic Inquiry* **30**, 509–542.
- Chierchia G** (1998) Reference to kinds across languages. *Natural Language Semantics* **6**, 339–405.
- Christophersen P** (1938) *The articles: A study of their theory and use in English*. Copenhagen: E. Munksgaard.
- Classen A** (2013) Multilingualism in the Middle Ages and the Early Modern age: the literary historical evidence. *Neophilologus* **97**, 131–145.
- Comrie B, Stone G and Polinsky M** (1996) *The Russian language in the twentieth century*. Oxford: Oxford University Press.
- Czypionka A and Kupisch T** (2019) (The) polar bears are pink. How (the) Germans interpret (the) definite articles in plural subject DPs. *Journal of Comparative Germanic Linguistics* **22**, 247–291.
- Dayal V** (2004) Number marking and (in)definiteness in kind terms. *Linguistics and Philosophy* **27**, 393–450.
- Demske U** (2001) *Merkmale und Relationen: Diachrone Studien zur Nominalphrase des Deutschen*. Berlin: De Gruyter.
- Demske U** (2020) The grammaticalization of the definite article in German: from demonstratives to weak definites. In Szczepaniak R and Flick J (eds), *Walking on the Grammaticalization Path of the Definite Article in German: Functional Main and Side Roads*. Amsterdam: Benjamins, pp. 43–73.
- Epstein R** (1995) The later stages in the development of the definite article: Evidence from French. In Andersen H (ed), *Papers from the 11th International Conference on Historical Linguistics 1993*. Amsterdam: John Benjamins, pp. 159–175.
- Fischer O, Norde M and Perridon H** (eds) (2004) *Up and down the cline: the nature of grammaticalization*. Amsterdam: John Benjamins.
- Fishman JA** (2001) 300-Plus Years of Heritage Language Education in the United States. In Payton JK, McGinnis D, Ranard A and S (eds), *Heritage Languages in America. Preserving a National Resource. Language in Education: Theory and Practice*. McHenry, IL: Delta Systems Inc, pp. 81–97.
- Flick J** (2021) *Die Entwicklung des Definitartikels im Althochdeutschen*. Berlin: Language Science Press.
- Flores C, Kupisch T and Rinke E** (2017) Linguistic Foundations of Heritage Language Development from the Perspective of Romance Languages in Germany. In Trifonas P and Aravossitas T (eds), *Handbook of Research and Practice in Heritage Language Education*. Champaign, IL: Springer, pp. 1–18. DOI: [https://doi.org/10.1007/978-3-319-38893-9\\_12-1](https://doi.org/10.1007/978-3-319-38893-9_12-1).
- Flores C and Rinke E** (2021) Portuguese as heritage language in Germany – a linguistic perspective. *Languages* **6**. DOI: <https://doi.org/10.3390/languages6010010>.
- Franco L, Manzini R and Savoia L** (2015) N morphology and its interpretation. The neuter in Central Italian varieties and its implications. *Isogloss* **14**, 41–67.
- Garzonio J and Poletto C** (2018) Exploiting microvariation: How to make the best of your incomplete data. *Glossa: A Journal of General Linguistics* **3**, 1–21.
- Givón T** (1979) *On understanding grammar*. Orlando, Florida: Academic Press.
- Givón T** (1981) On the development of the numeral *one* as an indefinite marker. *Folia Linguistica Historica* **2**, 35–53.
- Greenberg JH** (1978) How does a language acquire gender markers? In Greenberg JH, Ferguson CA and Moravcsik EA (eds), *Universals of Human Language* 3. Stanford CA: Stanford University Press, pp. 47–82.
- Grevisse M** (1964) *Le Bon Usage: Grammaire Française*. French & European Pubns.
- Gundel JK, Hedberg N and Zacharski R** (1993) Cognitive status and the form of referring expressions in discourse. *Language* **69**, 274–307.
- Haase M and Nau N** (1996) Einleitung: Sprachkontakt und Grammatikalisierung. *Sprachtypologie und Universalienforschung (STUF)* **49**, 3–8.
- Harves S** (2013) The genitive of negation in Russian. *Language and Linguistics Compass* **7**, 647–662.
- Haspelmath M** (2004) On directionality in language change with particular reference to grammaticalization. In Fischer O, Norde M and Perridon H (eds), *Up and down the cline: the nature of grammaticalization*. Amsterdam: John Benjamins, pp. 17–44.
- Heine B** (1997) *Cognitive Foundations of Grammar*. New York, Oxford: Oxford University Press.
- Heine B, Claudi U and Hünnemeyer F** (1991) *Grammaticalization: A conceptual framework*. Chicago: Chicago University Press.
- Heine B and Kuteva T** (2005) *Language contact and grammatical change*. Cambridge: CUP.
- Himmelmann N** (1997) *Deiktikon, Artikel, Nominalphrase: Zur Emergenz syntaktischer Struktur*. Tübingen: Max Niemeyer.
- Hopper P** (1987) Emergent Grammar. *Berkeley Linguistic Society* **13**, 139–157.
- Hopper P and Traugott EC** (1993) *Grammaticalization*. Cambridge: Cambridge University Press.
- Hopper P and Traugott EC** (2003) *Grammaticalization*. 2<sup>nd</sup> edition. Cambridge: Cambridge University Press.
- Jenks P** (2011) *The hidden structure of Thai noun phrases*. Ph.D. Dissertation, Harvard University.
- Kinn K** (2020) Stability and attrition in American Norwegian nominals: a view from predicate nouns. *The Journal of Comparative Germanic Linguistics* **23**, 3–38.
- Kirby S** (2017) Culture and biology in the origins of linguistic structure. *Psychonomic Bulletin and Review* **24**, 118–137.
- Kornfilt J** (1997) *Turkish*. New York: Routledge.
- Krifka M** (1995) Common nouns: A contrastive analysis of English and Chinese. In Carlson GN and Pelletier JF (eds), *The Generic Book*. Chicago: University of Chicago Press, pp. 398–411.
- Krifka M, Pelletier FJ, Carlson GN, ter Meulen A, Link G and Chierchia G** (1995) Genericity: An introduction. In Carlson GN and Pelletier JF (eds), *The Generic Book*. Chicago IL: Chicago University Press, pp. 1–124.
- Kupisch T** (2007) Determiners in bilingual German–Italian children: What they tell us about the relation between language influence and language dominance. *Bilingualism: Language and Cognition* **10**, 57–78.
- Kupisch T** (2012) Generic subjects in the Italian of early German–Italian bilinguals and German learners of Italian as a second language. *Bilingualism: Language and Cognition* **15**, 736–756.
- Kupisch T and Koops C** (2007) The definite article in non-specific noun phrases: comparing French and Italian. In Stark E, Leiss E and Abraham W (eds), *Nominal Determination. Typology, context constraints and historical emergence*. Amsterdam: John Benjamins, pp. 187–210.
- Kupisch T and Rothman J** (2018) Terminology Matters! Why Difference Is Not Incompleteness and How Early Child Bilinguals Are Heritage Speakers. *International Journal of Bilingualism* **22**, 564–582.
- Ledgeway A** (2012) From Latin to Romance: The Rise of Configurationality, Functional Categories and Head-marking. *Transactions of the Philological Society* **110**, 422–442.
- Lefebvre C, White L and Jourdan C** (2006) (eds) *L2 acquisition and creole genesis*. Amsterdam: John Benjamins.
- Lehmann C** (2015) *Thoughts on Grammaticalization: A programmatic sketch*. Berlin: Language Science Press.
- Leiss E** (2000) *Artikel und Aspekt: Die grammatischen Muster von Definitheit*. Berlin: Walter de Gruyter.
- Lightfoot DW** (1991) *How to set parameters: Arguments from language change*. Cambridge, MA: MIT Press.

- Lightfoot DW** (1999) *The development of language acquisition*. Oxford: Blackwell.
- Longobardi G** (1994) Reference and proper names: A theory of N-raising in syntax and logical form. *Linguistic Inquiry* 25, 609–665.
- Longobardi G** (1999) The structure of DPs: Some principles, parameters, and problems. In Baltin M and Collins C (eds), *The Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell, pp. 562–603.
- Longobardi G** (2011) How comparative is semantics. *Natural Language Semantics* 9, 335–369.
- Lyons C** (1999) *Definiteness*. Cambridge: Cambridge University Press.
- Meisel JM** (1986) Word order and case marking in early child language. Evidence from the simultaneous acquisition of two first languages: French and German. *Linguistics* 24, 123–183.
- Meisel JM** (2011a) *First and Second Language Acquisition: Parallels and Differences*. Cambridge: Cambridge University Press.
- Meisel JM** (2011b) Bilingual language acquisition and theories of diachronic change. *Bilingualism as cause and effect of grammatical change. Bilingualism: Language and Cognition* 14, 121–145.
- Meisel JM, Elsig M and Rinke E** (2013) *Language Acquisition and Change. A Morphosyntactic Perspective*. Edinburgh: Edinburgh University Press.
- Merlan FC** (1982) *Mangarayi*. Amsterdam: North-Holland.
- Michaelis SM and Haspelmath M** (2020) Grammaticalization in creole languages: Accelerated functionalization and semantic imitation. In Bisang W and Malchukov A (eds), *Grammaticalization scenarios: Cross-linguistic variation and universal tendencies*. Berlin: De Gruyter Mouton, pp. X–X.
- Montrul S** (2008) *Incomplete Acquisition in Bilingualism: Re-examining the Age Factor*. Amsterdam: John Benjamins.
- Montrul S** (2016) *The acquisition of heritage languages*. Cambridge: Cambridge University Press.
- Montrul S and Ionin T** (2012) Dominant language transfer in Spanish heritage speakers and L2 learners in the interpretation of definite articles. *The Modern Language Journal* 96, 70–94.
- Mustanoja T** (1960) *A Middle English Syntax*. Helsinki: Société néophilologique.
- Nagy N** (2016) Heritage languages as new dialects. In Côté MH, Knooihuizen R and Nerbonne J (eds), *The Future of Dialects: Selected Papers from Methods in Dialectology XV*. Berlin: Language Science, pp. 15–34.
- Nagy N** (2017) Cross-cultural approaches: Comparing heritage languages in Toronto. *Penn Working Papers in Linguistics* 23, 12.2.
- Nagy N and Lo S** (2019) Classifier use in Heritage and Hong Kong Cantonese. *Asia-Pacific Language Variation* 5, 84–108.
- Norde M** (2009) *Degrammaticalization*. Oxford: Oxford University Press.
- Oubouzar E** (1992) Zur Ausbildung des bestimmten Artikels im Althochdeutschen. In Desportes Y (ed), *Althochdeutsch. Syntax und Semantik*. Lyon: Université Lyon 1992, pp. 71–87.
- Oubouzar E** (1997) Zur Frage der Herausbildung eines bestimmten und eines unbestimmten Artikels im Althochdeutschen. *Cahiers d'études Germaniques* 32, 161–175.
- Pires A and Rothman J** (2009) Disentangling the sources of incomplete acquisition. *International Journal of Bilingualism* 13, 211–238.
- Polinsky M** (2018) *Heritage Language and their speakers*. Cambridge: Cambridge University Press.
- Polinsky M** (forthcoming) Syntactic properties of heritage languages. In Grohmann K (ed), *The Cambridge Handbook of Minimalism*.
- Polinsky M and Scontras G** (2020) Understanding Heritage Languages. *Bilingualism: Language and Cognition* 23, 4–20.
- Renzi L** (1976) Grammatica e storia dell'articolo italiano. *Studi di Grammatica Italiana* 5, 5–42.
- Rinke E and Flores C** (2014) Morphosyntactic knowledge of clitics by Portuguese heritage bilinguals. *Bilingualism: Language and Cognition* 17, 681–699.
- Ronai E** (2018) Quantifier scope in heritage bilinguals: a comparative experimental study. In Hucklebridge S and Nelson M (eds), *NELS 48: Proceedings of the 48th Annual Meeting of the North East Linguistic Society 3*. Amherst: GLSA, University of Massachusetts, pp. 29–38.
- Rothman J** (2007) Heritage speaker competence differences, language change, and input type: Inflected infinitives in Heritage Brazilian Portuguese. *International Journal of Bilingualism* 11, 359–389.
- Rothman J** (2009) Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages. *International Journal of Bilingualism* 13, 155–163.
- Saito M, Lin T.-H. J. and Murasugi K** (2008) N'-ellipsis and the structure of noun phrases in Chinese and Japanese. *Journal of East Asian Linguistics* 17, 247–271.
- Schmuck M** (2020) The rise of the onymic article in Early New High German: Areal factors and the triggering effect of bynames. In Szczepaniak R and Flick J (eds), *Walking on the grammaticalization path of the definite article*. Amsterdam: John Benjamins, pp. 200–226.
- Schmuck M and Szczepaniak R** (2014) Der Gebrauch des Definitartikels vor Familien- und Rufnamen im Frühneuhochdeutschen aus grammatikalisierungs-theoretischer Perspektive. In Debus F, Heuser R and Nübling D (eds), *Linguistik der Familiennamen*. Hildesheim: Olms, pp. 97–137.
- Schroeder C** (2006) Articles and article systems in some areas of Europe. In Bernini G (ed), *Pragmatic organization of discourse in the languages of Europe*. Berlin: De Gruyter, pp. 545–611.
- Scontras G, Polinsky M, Tsai C-YE and Mai K** (2017) Cross-linguistic scope ambiguity: When two systems meet. *Glossa: A Journal of General Linguistics* 2, 1–28.
- Serratrice L, Sorace A, Filiaci F and Baldo M** (2009) Bilingual Children's Sensitivity to Specificity and Genericity: Evidence from Metalinguistic Awareness. *Bilingualism: Language and Cognition* 12, 239–257.
- Sybesma RPE** (1992) *Causatives and accomplishments: The case of Chinese ba 1*. Holland Institute of Generative Linguistics.
- Szczepaniak R** (2011) *Grammatikalisierung im Deutschen: Einführung*. Tübingen: Narr.
- Szczepaniak R** (2016) Vom Zahlwort eins zum Indefinitartikel eine(e): Rekonstruktion des Grammatikalisierungsverlaufs im Alt- und Mittelhochdeutschen. In Bittner A and Köpcke K-M (eds), *Regularität und Irregularität in Phonologie und Morphologie: Diachron, kontrastiv, typologisch*. Berlin, Boston: De Gruyter, pp. 247–263.
- van Baal Y** (2020) *Compositional Definiteness in American Heritage Norwegian*. PhD dissertation, University of Oslo.
- von Fintel K** (1997) Bare plurals, bare conditionals, and only. *Journal of Semantics* 13, 1–56.
- Wei L and Lee S** (2001) L1 Development in an L2 Environment: The Use of Cantonese Classifiers and Quantifiers by Young British-Born Chinese in Tyneside. *International Journal of Bilingual Education and Bilingualism* 4, 359–382.
- Wiese H** (2013) What can new urban dialects tell us about internal language dynamics? The power of language diversity. *Linguistische Berichte* 19, 207–245.
- Wiese H and Rehbein I** (2015) Coherence in new urban dialects: a case study. *Lingua* 172/173, 45–61.
- Wilhelm A** (2008) Bare nouns and number in Dëne Sulinë. *Natural Language Semantics* 16, 39–68.
- Yang C** (2003) *Knowledge and Learning in Natural Language*. Oxford University Press.
- Yang C** (2016) *The Price of Linguistic Productivity*. MIT Press.